Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

Engineering cell identity: establishing new gene regulatory and chromatin landscapes

Chuner Guo[1,2,3] and Samantha A Morris[1,2,3]

Cellular reprogramming can be achieved by ectopically expressing transcription factors that directly convert one differentiated cell type into another, bypassing embryonic states. A number of different cell types have been generated by such 'direct lineage reprogramming' methods, but their practical utility has been limited because, in most protocols, the resulting populations are often partially differentiated or incompletely specified. Here, we review mechanisms of lineage reprogramming by pioneer transcription factors, a unique class of transcriptional regulators that has the capacity to engage with silent chromatin to activate target gene regulatory networks. We assess the possible barriers to successful reprogramming in the context of higher-order chromatin landscape, considering how the mechanistic relationship between nuclear organization and cell identity will be crucial to unlocking the full potential of cell fate engineering.

Addresses
[1] Department of Developmental Biology, Washington University School of Medicine in St. Louis, 660 S. Euclid Avenue, Campus Box 8103, St. Louis, MO 63110, USA
[2] Department of Genetics, Washington University School of Medicine in St. Louis, 660 S. Euclid Avenue, Campus Box 8103, St. Louis, MO 63110, USA
[3] Center of Regenerative Medicine, Washington University School of Medicine in St. Louis, 660 S. Euclid Avenue, Campus Box 8103, St. Louis, MO 63110, USA

Corresponding author: Morris, Samantha A (s.morris@wustl.edu)

Current Opinion in Genetics & Development 2017, **46**:50–57

This review comes from a themed issue on **Cell reprogramming**

Edited by **Jianlong Wang** and **Miguel Esteban**

http://dx.doi.org/10.1016/j.gde.2017.06.011

Introduction

We now know that, remarkably, a cell can be engineered to adopt another identity. The field has come a long way since Gurdon's 1962 landmark study on *Xenopus laevis* tadpoles derived by transplanting intestine epithelial nuclei into irradiated eggs [1], demonstrating that a terminally differentiated nucleus still possesses the capability to give rise to a full animal. More than half a century later, many differentiated cell types can now be engineered to become induced pluripotent stem cells (iPSCs), independent of an oocyte, using a variety of strategies such as transcription factor overexpression [2–4] and small molecule treatment [5,6] (reviewed extensively in [7–10]). These induced pluripotent stem cells can then be directed to differentiate into cells of various lineages (including cardiomyocytes [11–13], neurons [14], and pancreatic β cells [15,16]) for patient-specific disease modeling, drug screening, and therapy (reviewed in [17,18]). An alternate strategy aims to convert cell identity between two fully-differentiated states, without reverting to the pluripotent state. One of the trailblazer studies of this so-called 'direct lineage reprogramming' was the MyoD-mediated induction of muscle fate in various differentiated primary cells and cell lines [19]. It is now possible to generate cell types of all three germ layers (e.g., induced cardiomyocytes [20], induced neurons [21], induced pancreatic beta cells [22]) from fully differentiated cell types, expanding the arsenal of strategies available for cell fate engineering.

Engineering cell identity: are we there yet?

As new strategies for cell and tissue engineering arise, key questions remain in the field. One question pertaining to direct lineage reprogramming is whether the process does indeed directly reprogram cell identity between fully differentiated states. Although many published studies report no intermediate states on the basis of lineage tracing experiments that monitored canonical progenitor gene expression [20,22], it is possible that non-canonical intermediates are visited during the process. This possibility is logical, given that many developmental gene regulatory networks (GRNs) are engaged by the transcription factors used in reprogramming, most of which are known to be developmentally important genes themselves [23]. Experimentally, this is supported by recent observations in neurons reprogrammed directly from fibroblasts, where a fractional identity was observed mid-conversion that was not accounted for by fibroblast, neuron, or neuronal progenitor identities based on analysis of single-cell RNA-sequencing data [24••]. Another example is illustrated in our recent study of cell type classification by our network biology platform, CellNet, which revealed that the transcription factor combination previously thought to generate induced hepatocytes [25] in fact generated immature endoderm progenitor cells with broad developmental potential as well as long-term self-renewal [26].

Another key issue to be addressed is the low efficiency of cell identity engineering. Many protocols are fraught with

low reprogramming efficiencies, estimated to range from 0.1% [2] to 20% [21]. Direct lineage reprogramming, although not without its own advantages as mentioned above, suffers especially because many starting cell types do not self-renew indefinitely and are therefore difficult to scale up. This poses a barrier to useful downstream applications such as cell replacement therapy, which often requires numbers of cells on the order of millions, and can be especially problematic for the generation of post-mitotic cell types that do not passage via a proliferative intermediate or only transit through proliferation briefly.

Finally, at the heart of the discussion is the question of whether the engineered cell has completely adopted the desired identity. Although functionally many engineered cells appear to mimic their target cell types, close inspection of their transcriptomes and epigenomes reveals that cell identities are oftentimes insufficiently specified, and that starting cell identities often fail to be fully erased [26,27]. In the case of induced hepatocytes generated by overexpression of *Hnf4a* and either *Foxa1*, *2*, or *3* [25], the converted cells failed to classify as liver, and were found to possess intestinal signatures, as well as significant remaining starting fibroblast signatures [26]. This observed 'infidelity' of many reprogramming strategies may be at least in part because of unrefined reprogramming 'recipes', or incomplete transcription factor cocktails that fail to engage all GRNs needed to fully establish identity. The starting cell identity remains perhaps because canonical developmental GRNs, while engaged and activated by reprogramming factors, may be unable to execute full erasure of the starting cell identity. Ultimately, to perfectly and precisely engineer cell identities at a desirable efficiency, it will be important to understand the mechanism by which this is achieved in the context of gene regulation, that is, how to turn on appropriate genetic programs and turn off irrelevant ones, and also importantly, in the context of nuclear organization, that is, how to reposition the respective genetic programs amidst current chromatin landscapes.

Pioneer transcription factors engage silent chromatin to reprogram cell identity

Many reprogramming strategies include the use of one or more transcription factors or small molecules to activate core GRNs. A common strategy for identifying the best 'recipe' is to simultaneously overexpress a pool of transcription factors known to be of developmental significance, then subtract one by one from the pool until the minimal set is determined [2,20–22]. Many of these transcription factor recipes include pioneer transcription factors such as OCT4, SOX2, and KLF4 in the case of iPSCs [2,3], ASCL1 for induced neurons [21], GATA4 for induced cardiomyocytes [20], and FOXA1/2/3 for induced hepatocytes [25] (Table 1). In addition to binding to DNA, pioneer transcription factors are known to have the capability to bind compacted chromatin, increase target site accessibility, and recruit cooperative transcription factors to bind (Figure 1, and reviewed in [23,28,29]). This initial binding event by pioneer transcription factors often precedes transcriptional activation during development, and is thought to be important for establishing competence for additional regulatory events and developmental programs.

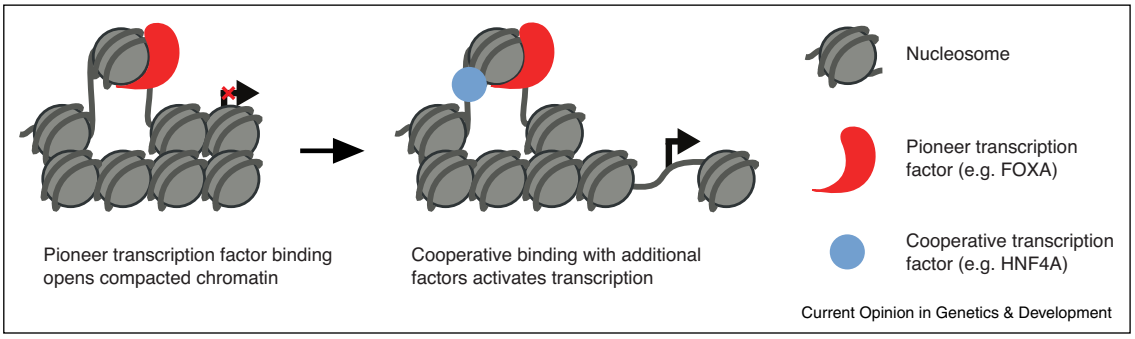
The FOXA family represents the first pioneer transcription factors illustrated to play important roles in regulating lineage specification during development. Developmentally, *Foxa* is widely expressed during embryogenesis starting around the time of gastrulation in the mouse, most prominently within the endoderm and its derivatives such as lung, liver, pancreas, and intestine [31]. Remarkably, despite its broad expression pattern during endoderm development, FOXA specifies cell fates in a context-dependent manner. In liver precursor cells,

Table 1

Selected pioneer transcription factors used in reprogramming fibroblasts to other cell types

Pioneer transcription factor	Cooperative transcription factor	Target cell identity
OCT4, SOX2, KLF4	MYC	iPSC [3]
OCT4	RUNX2, OSX, MYCL	Osteoblast [61]
C/EPBA, PU.1	–	Macrophage [62]
OCT4	–	Hematopoietic progenitor [63]
GATA2	GFI1B, FOS, ETV6	Hematopoietic progenitor [64]
GATA4	MEF2C, TBX5	Cardiomyocyte [20]
FOXA1/2/3	HNF4A	Hepatocyte [25]
		Endoderm progenitor [26]
FOXA3, GATA4	HNF1A	Hepatocyte [65]
ASCL1	BRN2, MYT1L, LHX3, HB9, ISL1, NGN2	Motor neuron [66]
ASCL1, FOXA2	BRN2, MYT1L, LMX1A	Dopaminergic neuron [67]
ASCL1	LMX1A, NURR1	Dopaminergic neuron [68]
ASCL1	BRN2, MYT1L, (NEUROD1)	Glutamatergic neuron [21,69]
ASCL1	MYT1L, NEUROD1/2, miR-9/9*, miR-124	Glutamatergic neuron [70]
SOX2	BRN2, FOXG1	Neural progenitor [71]

Figure 1



Pioneer transcription factor primes compacted chromatin to establish competence. Pioneer transcription factors can displace linker histones to bind and open locally compacted nucleosomal chromatin. This priming action increases binding site accessibility and establishes competence for lineage-specific transcriptional activation, which requires cooperative binding of additional transcription factors.

FOXA is among the first to bind to the silent albumin (Alb) enhancer, preceding other enhancer-binding factors and transcriptional activation [30]. In the prospective intestinal endoderm, FOXA expression equips cells with hepatic potential, which is normally inhibited by the overlying intestinal mesoderm [31]. Indeed, FOXA binding is shown to induce a 'poised' chromatin state, priming the enhancers for lineage-specific recruitment of additional transcription factors and activation of the respective genetic programs [32].

This pioneer activity of FOXA is characterized by its ability to interact with core histone proteins, displace linker histones, and open compacted chromatin [30]. Structurally, FOXA's *forkhead* domain resembles the globular domain of avian erythrocyte linker histone H5, a member of the linker histone H1 family [33]. The ability of FOXA to engage compacted chromatin is key to its pioneering action, which endows competence for cooperative binding of additional factors, and is required for maintaining accessible nucleosomes for lineage-specific gene activation [34•].

Higher order chromatin landscape may impose a barrier to establishing new identity

Although pioneer transcription factors play a central role in engaging compacted chromatin and establishing competence, they alone do not appear sufficient to fully convert cells to a new identity. Even together with other transcription factors determined to be the best 'recipe' for reprogramming, cell identities are often incompletely specified, often at very low efficiencies. This could be explained by the following:

(1) Pioneer transcription factors bind compacted chromatin promiscuously to establish developmental potential, but additional signals and cues are needed to precisely guide the cells down developmental trajectories to become fully mature. This is supported by the unexpected intestinal fate in induced hepatocytes generated by *Hnf4a* and *Foxa1/2/3* overexpression [26], as well as the unexpected myogenic fate arising from neuronal reprogramming using the well-established 'on-target' pioneer transcription factor ASCL1, in the absence of maturation factors MYT1L and BRN2 [24••,35,36]. Furthermore, direct lineage reprogramming strategies that are carried out *in vivo* or involve transplantation of engineered cells often show better specification and maturation of target cell identities [22,37]. This is at least in part because of the possibility that the *in vivo* environment provides additional signals and cues to aid maturation.

(2) Highly repressive chromatin conformation is restrictive to pioneer transcription factor binding and precludes access to required genetic programs. For example, megabase heterochromatin domains marked by H3K9me3 in fibroblasts contain genes required for reprogramming to pluripotency but are initially resistant to binding by OSKM (OCT4, SOX2, KLF4, MYC) factors [38]. This is in agreement with the hypothesis that high order compaction of chromatin may be largely inaccessible even for pioneer transcription factors. Indeed, upon transient knockdown of related histone methyltransferases in fibroblasts, initial binding by OSKM was significantly increased [38]. Furthermore, reprogramming to pluripotency possesses an early stochastic phase and a late deterministic phase [39], supporting the idea that initial access to repressive chromatin compartments remains a barrier to entering deterministic reprogramming. In line with this hypothesis, there is abundant evidence that chromatin epigenetics can be modulated by small molecules, often to increase reprogramming efficiencies [40,41] (reviewed in [42,43]). Certain chemicals improve reprogramming by regulating histone modifications, DNA methylation, or chromatin remodeling

complexes [44], which may be synergizing to the action of transcription factors.

Clearly then, genome accessibility is deeply relevant to gene regulation and the control of cell fate. Consider cell fate reprogramming in the context of the classic Waddington's epigenetic landscape model of development, where a cell must traverse uphill from a low valley in order to dedifferentiate to pluripotency or transdifferentiate to another cell identity (reviewed in [45]). This 'hill' then acts as an epigenetic barrier that must be overcome when cell fate changes are being driven by exogenous means. In the context of engineering cell identity, the higher-order architecture of nuclear organization is likely a component of the epigenetic barrier, where large-scale structures such as the megabase heterochromatin domains prevent exogenous factors from swiftly inducing epigenetic remodeling and transcriptional shifts. In the text below, we will discuss the higher-order nuclear organization systematically in the context of epigenetic barriers and strategies to overcome them for cell identity engineering.

Nuclear organization is a meaningful component of cell identity

Consider the challenge of packing the entire length of the genome (2 m) into a nucleus with the diameter of typically a few microns — the equivalent of 'packing 40 km of fine yarn into a tennis ball' [46]. One can imagine that orderly packaging of genetic material and segregation of the packaged chromatin into different spatial compartments must be crucial for establishing proper control of gene regulation. Recent advances in chromosome conformation capture (3C) and other C-technologies have allowed for increasingly high resolutions at which we can examine genomic interactions, as well as increasingly high throughputs so that multiple tissue and cell types can be studied simultaneously. Indeed, studies over the past few decades have provided insights into nuclear organization across various cell types (reviewed in [47,48]).

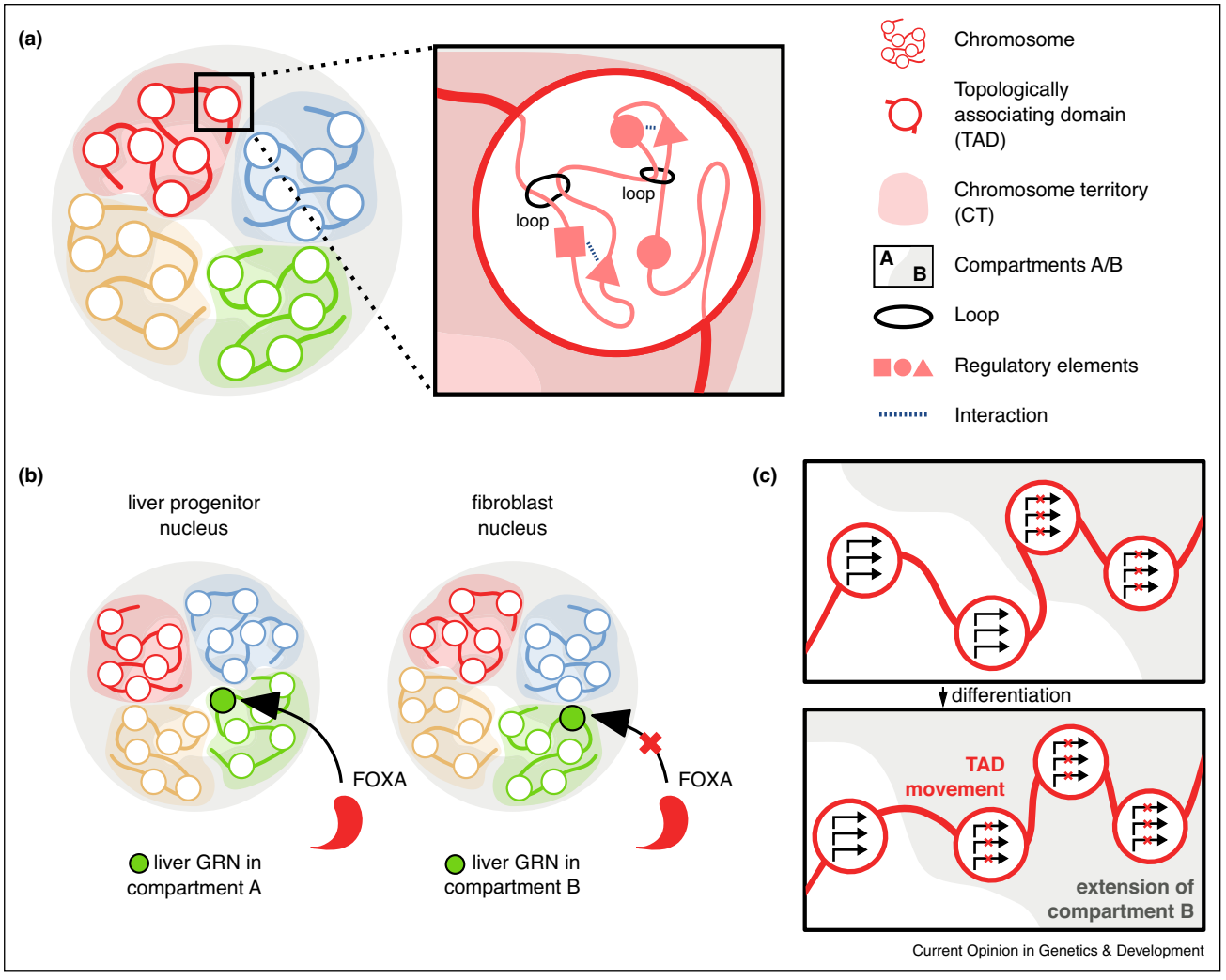
A bird's eye view of the chromatin landscape reveals that the nucleus is organized hierarchically (Figure 2a). At the microscopic level, individual chromosomes occupy distinctive chromosome territories (CTs) [49]. This was confirmed by Hi-C experiments showing that intrachromosomal loci interacted at a much higher frequency than interchromosomal loci [50]. On a subchromosomal level, we have known that active euchromatin and inactive heterochromatin spatially segregate into different regions; these regions were later computationally confirmed by spatial compartmentalization of Hi-C analysis, and were then termed compartments A and B [50]. The significance of compartmentalization dynamics to cell identity is illustrated in a study of subchromosomal compartments across different cell types, where reorganization on the sub-megabase scale was observed during differentiation of mouse embryonic stem cells (ESCs) to neuronal progenitor cells [51]. Extensive A/B compartment switching altering up to 36% of the compartments was also shown in a similar experiment using Hi-C to interrogate genomic interactions in human ESCs and four ESC-derived lineages, where an extension of the B compartment and lineage-specific compartment transitions were observed [52••]. As mentioned previously, the repressive chromatin in compartment B may be inaccessible to pioneer transcription factors such as FOXA, blocking access to targets that may be needed to establish target cell identity (Figure 2b). Furthermore, the expansion of compartment B during differentiation may explain the relative inefficiency of direct lineage reprogramming from differentiated cell types, compared to that of a directed differentiation approach from a pluripotent state.

Locally, nuclear organization occurs on a scale that facilitates regulation in a cell identity-specific manner (Figure 2a). Chromatin self-interacts to form topologically associating domains (TADs) with the help of architectural proteins such as CTCF and cohesin [53]. Within TADs, chromatin is further packaged with the aid of architectural proteins, resulting in local structures termed chromatin loops [54•]. Although TAD boundaries are largely conserved [53], TAD positioning and local structures within TADs can be dynamic across cell and tissue types [55], as well as over developmental processes. During stem cell differentiation, compartment switching occurs and results in expansion of compartment B, reflecting the movement of TADs across compartments (Figure 2c) [48,52••]. In the reverse process of reprogramming somatic cells to pluripotency, TADs and local structures undergo drastic reorganization, leading to the erasure of tissue-specific compartment configuration and establishment of ESC-like topology; however, early-passage iPSCs still possess cell origin-dependent structures acquired during reprogramming that do not represent remnants of somatic memory, alluding to the possibly distinct reprogramming trajectories that are cell origin-dependent [56••]. Notably, a more detailed view of reprogramming to pluripotency focused on regions of key genes revealed local topologies susceptible to incomplete rewiring, leading to aberrant gene expression, both of which can be repaired under 2i/LIF culture conditions [57••]. These two studies of different resolutions highlight the many levels of nuclear reorganization that occurs during reprogramming. Further research will shed light on conditions that may improve the precision of cell identity engineering by ensuring the establishment of the correct chromatin landscape.

Thoughts and perspectives

In summary, current evidence suggests that nuclear organization is non-random. Tissue-specific and cell type-specific interactions occur on various scales, and reorganization of the nucleus takes place during development. We also know that cell identities are most accurately

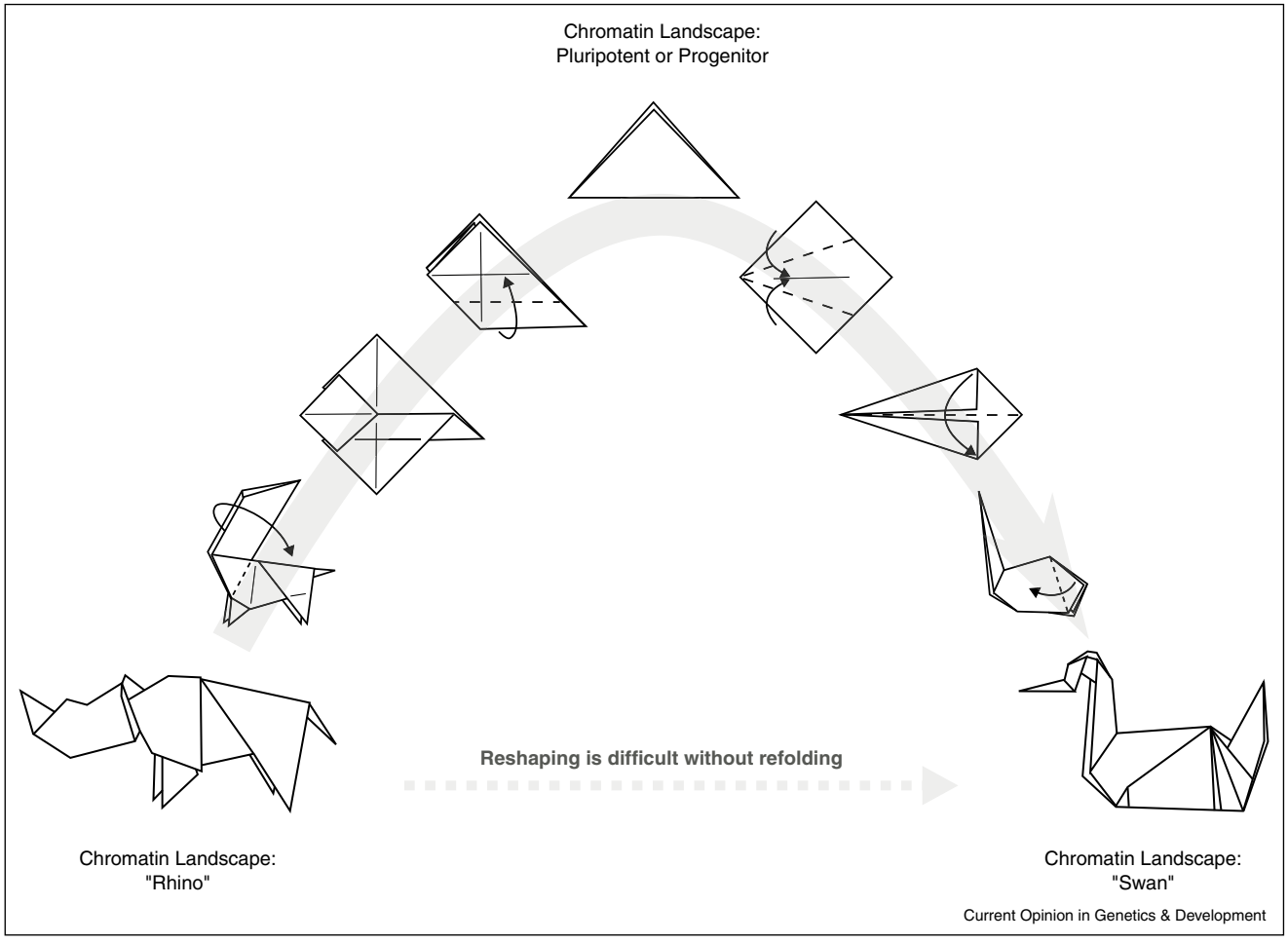
Figure 2



The nucleus is organized hierarchically. **(a)** On the chromosomal level, individual chromosomes are found within distinct chromosome territories (CTs). Active euchromatin and inactive heterochromatin segregate into subchromosomal compartments A and B. Topologically associating domains (TADs) are seen in both compartments A and B. TADs are structural units containing sequences that frequently interact with other sequences in the same TAD. Within TADs are local structures such as loops, which facilitate contact between gene regulatory elements. **(b)** Pioneer transcription factors may be incapable of accessing chromatin within compartment B. For example, FOXA may access relevant developmental GRNs within TADs in liver progenitor cells, where these TADs are positioned within compartment A. In a fully differentiated cell type of a different lineage such as fibroblasts, FOXA may be restricted from these GRNs that now reside in compartment B. **(c)** Nuclear organization is dynamic over developmental processes. TADs can move across compartments during differentiation, accompanied by the expansion of compartment B. Gene expression may be downregulated as a TAD moves into compartment B.

engineered when guided down the developmental trajectories by directed differentiation, and that it is often easier to directly reprogram immature cell types such as embryonic or neonatal fibroblasts compared to adult fibroblasts. Taken together, we propose that hierarchical refolding of the nucleus is required for a cell to establish a new identity, and that the best way to do so may be to return to a precursor state of chromatin organization, that is, a progenitor or pluripotent state in development, where a large part of the genome remains accessible to restructuring. It may be infeasible to shut down one terminally differentiated identity and create a new one in the same process because of the inaccessibility of target genetic programs that now reside in compartment B. Take the analogy of origami: once folded, an origami animal cannot simply be molded to adopt another shape, unless one unfolds it toward earlier stages (Figure 3). Further examination of the chromatin landscapes of partially or fully engineered cells may reveal roadblocks that are currently limiting the action of reprogramming factors, as well as targets that can be perturbed to modulate reprogramming efficiency and fidelity.

Figure 3



Establishment of a new cell identity may require a hierarchical reorganization of the chromatin landscape. Origami, the art of paper folding, is analogous to nuclear organization in the context of establishing cell identities. The chromatin landscape of one cell is different from another's ('rhino' vs. 'swan'). To convert a paper rhino into a swan, it is necessary to unfold the rhino such that the paper can be properly refolded into the shape of the swan, as direct reshaping may be challenging. Analogously, hierarchical reorganization of the nucleus and transit through an organizational precursor (corresponding to pluripotency or a progenitor state) may be necessary to ensure proper reprogramming of cell identities.

It should be noted that many of the observations from studying nuclear organization do not yet reveal the driving force behind the association between genomic elements, and its mechanistic relationship with the ultimate functional consequence of gene regulation, that is, the biology of a cell. As a recent review on single-cell genomic technologies points out [58], we now have an unprecedented array of tools to turn observations into insights, to translate phenomenology into a mechanistic understanding of fundamental biology. The resolution afforded by single-cell technologies also brings forth questions regarding the potential heterogeneity of cells sharing the same identity. Indeed, recent single-cell Hi-C studies reveal a stochasticity in local chromatin organization [59,60]. For us cell engineers, it will be important to elucidate the mechanism responsible for the initiation and execution of chromatin landscape remodeling that occurs during normal development, which could provide an actionable toolset to achieve precise and efficient cell identity engineering.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Gurdon JB: **The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles**. *J Embryol Exp Morphol* 1962, **10**:622-640.

2. Takahashi K, Yamanaka S: **Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors**. *Cell* 2006, **126**:663-676.

3. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S: **Induction of pluripotent stem cells from adult human fibroblasts by defined factors**. *Cell* 2007, **131**:861-872.

4. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R *et al.*: **Induced pluripotent stem cell lines derived from human somatic cells**. *Science* 2007, **318**:1917-1920.

5. Hou P, Li Y, Zhang X, Liu C, Guan J, Li H, Zhao T, Ye J, Yang W, Liu K *et al.*: **Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds**. *Science* 2013, **341**:651-654.

6. Zhu S, Li W, Zhou H, Wei W, Ambasudhan R, Lin T, Kim J, Zhang K, Ding S: **Reprogramming of human primary somatic cells by OCT4 and chemical compounds**. *Cell Stem Cell* 2010, **7**:651-655.

7. González F, Boué S, Izpisúa Belmonte JC: **Methods for making induced pluripotent stem cells: reprogramming à la carte**. *Nat Rev Genet* 2011, **12**:231-242.

8. Zhang Y, Li W, Laurent T, Ding S: **Small molecules, big roles — the chemical manipulation of stem cell fate and somatic cell reprogramming**. *J Cell Sci* 2012, **125**:5609-5620.

9. Yamanaka S: **Induced pluripotent stem cells: past, present, and future**. *Cell Stem Cell* 2012, **10**:678-684.

10. Takahashi K, Yamanaka S: **A decade of transcription factor-mediated reprogramming to pluripotency**. *Nat Rev Mol Cell Biol* 2016, **17**:183-193.

11. Carvajal-Vergara X, Sevilla A, Souza SLD', Ang Y-S, Schaniel C, Lee D-F, Yang L, Kaplan AD, Adler ED, Rozov R *et al.*: **Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome**. *Nature* 2010, **465**:808-812.

12. Burridge PW, Thompson S, Millrod MA, Weinberg S, Yuan X, Peters A, Mahairaki V, Koliatsos VE, Tung L, Zambidis ET: **A universal system for highly efficient cardiac differentiation of human induced pluripotent stem cells that eliminates interline variability**. *PLoS ONE* 2011, **6**:e18293.

13. Lian X, Hsiao C, Wilson G, Zhu K, Hazeltine LB, Azarin SM, Raval KK, Zhang J, Kamp TJ, Palecek SP: **Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling**. *Proc Natl Acad Sci U S A* 2012, **109**:E1848-E1857.

14. Dimos JT, Rodolfa KT, Niakan KK, Weisenthal LM, Mitsumoto H, Chung W, Croft GF, Saphier G, Leibel R, Goland R *et al.*: **Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons**. *Science* 2008, **321**:1218-1221.

15. Zhang D, Jiang W, Liu M, Sui X, Yin X, Chen S, Shi Y, Deng H: **Highly efficient differentiation of human ES cells and iPS cells into mature pancreatic insulin-producing cells**. *Cell Res* 2009, **19**:429-438.

16. Pagliuca FW, Millman JR, Gürtler M, Segel M, Van Dervort A, Ryu JH, Peterson QP, Greiner D, Melton DA: **Generation of functional human pancreatic β cells in vitro**. *Cell* 2014, **159**:428-439.

17. Avior Y, Sagi I, Benvenisty N: **Pluripotent stem cells in disease modelling and drug discovery**. *Nat Rev Mol Cell Biol* 2016, **17**:170-182.

18. Trounson A, DeWitt ND: **Pluripotent stem cells progressing to the clinic**. *Nat Rev Mol Cell Biol* 2016, **17**:194-200.

19. Davis RL, Weintraub H, Lassar AB: **Expression of a single transfected cDNA converts fibroblasts to myoblasts**. *Cell* 1987, **51**:987-1000.

20. Ieda M, Fu J-D, Delgado-Olguin P, Vedantham V, Hayashi Y, Bruneau BG, Srivastava D: **Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors**. *Cell* 2010, **142**:375-386.

21. 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**:1035-1041.

22. Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA: **In vivo reprogramming of adult pancreatic exocrine cells to beta-cells**. *Nature* 2008, **455**:627-632.

23. Morris SA: **Direct lineage reprogramming via pioneer factors; a detour through developmental gene regulatory networks**. *Development* 2016, **143**:2696-2705.

24. •• Treutlein B, Lee QY, Camp JG, Mall M, Koh W, Shariati SAM, Sim S, Neff NF, Skotheim JM, Wernig M *et al.*: **Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq**. *Nature* 2016, **534**:391-395.
Using single-cell RNA-seq, the authors studied the transcriptomics of direct lineage reprogramming from fibroblasts to induced neurons at multiple time points. The study revealed an unexpected alternative myogenic fate when the well-established proneural pioneer factor Ascl1 was used alone, as well as a reprogramming intermediate that was transcriptionally distinct from starting and target cell programs.

25. Sekiya S, Suzuki A: **Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors**. *Nature* 2011, **475**:390-393.

26. Morris SA, Cahan P, Li H, Zhao AM, San Roman AK, Shivdasani RA, Collins JJ, Daley GQ: **Dissecting engineered cell types and enhancing cell fate conversion via CellNet**. *Cell* 2014, **158**:889-902.

27. Cahan P, Li H, Morris SA, Lummertz da Rocha E, Daley GQ, Collins JJ: **CellNet: network biology applied to stem cell engineering**. *Cell* 2014, **158**:903-915.

28. Zaret KS, Carroll JS: **Pioneer transcription factors: establishing competence for gene expression**. *Genes Dev* 2011, **25**:2227-2241.

29. Iwafuchi-Doi M, Zaret KS: **Pioneer transcription factors in cell reprogramming**. *Genes Dev* 2014, **28**:2679-2692.

30. Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS: **Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4**. *Mol Cell* 2002, **9**:279-289.

31. Bossard P, Zaret KS: **Repressive and restrictive mesodermal interactions with gut endoderm: possible relation to Meckel's Diverticulum**. *Development* 2000, **127**:4915-4923.

32. Wang A, Yue F, Li Y, Xie R, Harper T, Patel NA, Muth K, Palmer J, Qiu Y, Wang J *et al.*: **Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates**. *Cell Stem Cell* 2015, **16**:386-399.

33. Clark KL, Halay ED, Lai E, Burley SK: **Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5**. *Nature* 1993, **364**:412-420.

34. • Iwafuchi-Doi M, Donahue G, Kakumanu A, Watts JA, Mahony S, Pugh BF, Lee D, Kaestner KH, Zaret KS: **The pioneer transcription factor FoxA maintains an accessible nucleosome configuration at enhancers for tissue-specific gene activation**. *Mol Cell* 2016, **62**:79-91.
This study further demonstrated that pioneer transcription factor, FoxA can open compacted nucleosomes and establish competence for transcriptional activation in a tissue-specific manner. The authors also showed that FoxA is required to maintain an accessible nucleosome configuration.

35. Wapinski OL, Vierbuchen T, Qu K, Lee QY, Chanda S, Fuentes DR, Giresi PG, Ng YH, Marro S, Neff NF *et al.*: **Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons**. *Cell* 2013, **155**:621-635.

36. Chanda S, Ang CE, Davila J, Pak C, Mall M, Lee QY, Ahlenius H, Jung SW, Südhof TC, Wernig M: **Generation of induced neuronal cells by the single reprogramming factor ASCL1**. *Stem Cell Reports* 2014, **3**:282-296.

37. Qian L, Huang Y, Spencer CI, Foley A, Vedantham V, Liu L, Conway SJ, Fu J, Srivastava D: **In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes**. *Nature* 2012, **485**:593-598.

38. Soufi A, Donahue G, Zaret KS: **Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome**. *Cell* 2012, **151**:994-1004.

39. Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, Klemm SL, van Oudenaarden A, Jaenisch R: **Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase**. *Cell* 2012, **150**:1209-1222.

40. Zhou Y, Wang L, Vaseghi HR, Liu Z, Lu R, Alimohamadi S, Yin C, Fu J-D, Wang GG, Liu J *et al.*: **Bmi1 is a key epigenetic barrier to direct cardiac reprogramming**. *Cell Stem Cell* 2016, **18**:382-395.

41. Smith DK, Yang J, Liu M-L, Zhang C-L: **Small molecules modulate chromatin accessibility to promote NEUROG2-mediated fibroblast-to-neuron reprogramming**. *Stem Cell Reports* 2016, **7**:955-969.

42. Soufi A: **Mechanisms for enhancing cellular reprogramming**. *Curr Opin Genet Dev* 2014, **25**:101-109.

43. Xu J, Du Y, Deng H: **Direct lineage reprogramming: strategies, mechanisms, and applications**. *Cell Stem Cell* 2015, **16**:119-134.

44. Yu C, Liu K, Tang S, Ding S: **Chemical approaches to cell reprogramming**. *Curr Opin Genet Dev* 2014, **28**:50-56.

45. Takahashi K: **Cellular reprogramming — lowering gravity on Waddington's epigenetic landscape**. *J Cell Sci* 2012, **125**:2553-2560.

46. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P: **Chromosomal DNA and its packaging in the chromatin fiber [Internet]**. *Molecular Biology of the Cell*. 2002.
Garland Science

47. Bickmore WA, van Steensel B: **Genome architecture: domain organization of interphase chromosomes**. *Cell* 2013, **152**:1270-1284.

48. Gibcus JH, Dekker J: **The hierarchy of the 3D genome**. *Mol Cell* 2013, **49**:773-782.

49. Cremer T, Cremer C: **Chromosome territories, nuclear architecture and gene regulation in mammalian cells**. *Nat Rev Genet* 2001, **2**:292-301.

50. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO *et al.*: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome**. *Science* 2009, **326**:289-293.

51. Phillips-Cremins JE, Sauria MEG, Sanyal A, Gerasimova TI, Lajoie BR, Bell JSK, Ong C-T, Hookway TA, Guo C, Sun Y *et al.*: **Architectural protein subclasses shape 3D organization of genomes during lineage commitment**. *Cell* 2013, **153**:1281-1295.

52. •• Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W *et al.*: **Chromatin architecture reorganization during stem cell differentiation**. *Nature* 2015, **518**:331-336.
This study showed that higher-order chromatin organization undergoes dynamic changes during differentiation. The authors used Hi-C to demonstrate that extensive compartment switching occurs in a lineage-specific manner during the differentiation of human ESCs to four different lineages, and that this likely reflects the movement of single or multiple TADs across compartments.

53. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions**. *Nature* 2012, **485**:376-380.

54. • Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES *et al.*: **A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping**. *Cell* 2014, **159**:1665-1680.
This recent study further demonstrated that chromatin architecture is organized in a cell type-specific and tissue-specific manner by analyzing local interactions at a high resolution using Hi-C data from multiple cell and tissue types.

55. Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, Li Y, Lin S, Lin Y, Barr CL *et al.*: **A compendium of chromatin contact maps reveals spatially active regions in the human genome**. *Cell Rep* 2016, **17**:2042-2059.

56. •• Krijger PHL, Di Stefano B, de Wit E, Limone F, van Oevelen C, de Laat W, Graf T: **Cell-of-origin-specific 3D genome structure acquired during somatic cell reprogramming**. *Cell Stem Cell* 2016, **18**:597-610.
This study showed that chromatin reorganization is observed in reprogramming somatic cells to pluripotency, in effect resetting the chromatin landscape. However, cell origin-specific features can still be seen in early-passage iPSCs, suggesting that these may be acquired during distinct reprogramming trajectories from different somatic cell types.

57. •• Beagan JA, Gilgenast TG, Kim J, Plona Z, Norton HK, Hu G, Hsu SC, Shields EJ, Lyu X, Apostolou E *et al.*: **Local genome topology can exhibit an incompletely rewired 3D-folding state during somatic cell reprogramming**. *Cell Stem Cell* 2016, **18**:611-624.
This study took an approach similar to that of Ref. [56••], but with a higher resolution focusing on a few key gene regions, to look at local chromatin changes in mouse ESCs, primary neural progenitor cells (NPCs), and NPC-derived iPSCs. The authors found that local chromatin organization can be incompletely rewired in reprogrammed iPSCs, leading to aberrant gene expression levels, and that the 2i/LIF culturing condition can promote proper wiring of chromatin interactions as well as accurate gene expression in the reprogrammed cells.

58. Tanay A, Regev A: **Scaling single-cell genomics from phenomenology to mechanism**. *Nature* 2017, **541**:331-338.

59. Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, Leeb M, Wohlfahrt KJ, Boucher W, O'Shaughnessy-Kirwan A *et al.*: **3D structures of individual mammalian genomes studied by single-cell Hi-C**. *Nature* 2017, **544**:59-64.

60. Flyamer IM, Gassler J, Imakaev M, Brandão HB, Ulianov SV, Abdennur N, Razin SV, Mirny LA, Tachibana-Konwalski K: **Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition**. *Nature* 2017, **544**:110-114.

61. Yamamoto K, Kishida T, Sato Y, Nishioka K, Ejima A, Fujiwara H, Kubo T, Yamamoto T, Kanamura N, Mazda O: **Direct conversion of human fibroblasts into functional osteoblasts by defined factors**. *Proc Natl Acad Sci U S A* 2015, **112**:6152-6157.

62. Feng R, Desbordes SC, Xie H, Tillo ES, Pixley F, Stanley ER, Graf T: **PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells**. *Proc Natl Acad Sci U S A* 2008, **105**:6057-6062.

63. Szabo E, Rampalli S, Risueño RM, Schnerch A, Mitchell R, Fiebig-Comyn A, Levadoux-Martin M, Bhatia M: **Direct conversion of human fibroblasts to multilineage blood progenitors**. *Nature* 2010, **468**:521-526.

64. Pereira C-F, Chang B, Qiu J, Niu X, Papatsenko D, Hendry CE, Clark NR, Nomura-Kitabayashi A, Kovacic JC, Ma'ayan A *et al.*: **Induction of a hemogenic program in mouse fibroblasts**. *Cell Stem Cell* 2013, **13**:205-218.

65. Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, Hu Y, Wang X, Hui L: **Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors**. *Nature* 2011, **475**:386-389.

66. Son EY, Ichida JK, Wainger BJ, Toma JS, Rafuse VF, Woolf CJ, Eggan K: **Conversion of mouse and human fibroblasts into functional spinal motor neurons**. *Cell Stem Cell* 2011, **9**:205-218.

67. Pfisterer U, Kirkeby A, Torper O, Wood J, Nelander J, Dufour A, Björklund A, Lindvall O, Jakobsson J, Parmar M: **Direct conversion of human fibroblasts to dopaminergic neurons**. *Proc Natl Acad Sci U S A* 2011, **108**:10343-10348.

68. Caiazzo M, Dell'Anno MT, Dvoretskova E, Lazarevic D, Taverna S, Leo D, Sotnikova TD, Menegon A, Roncaglia P, Colciago G *et al.*: **Direct generation of functional dopaminergic neurons from mouse and human fibroblasts**. *Nature* 2011, **476**:224-227.

69. Pang ZP, Yang N, Vierbuchen T, Ostermeier A, Fuentes DR, Yang TQ, Citri A, Sebastiano V, Marro S, Südhof TC *et al.*: **Induction of human neuronal cells by defined transcription factors**. *Nature* 2011, **476**:220-223.

70. Yoo AS, Sun AX, Li L, Shcheglovitov A, Portmann T, Li Y, Lee-Messer C, Dolmetsch RE, Tsien RW, Crabtree GR: **MicroRNA-mediated conversion of human fibroblasts to neurons**. *Nature* 2011, **476**:228-231.

71. Lujan E, Chanda S, Ahlenius H, Südhof TC, Wernig M: **Direct conversion of mouse fibroblasts to self-renewing, tripotent neural precursor cells**. *Proc Natl Acad Sci U S A* 2012, **109**:2527-2532.