npg

A blueprint for engineering cell fate: current technologies to reprogram cell identity

Samantha A Morris[1, 2, 3], George Q Daley[1, 2, 3]

[1]Stem Cell Transplantation Program, Division of Pediatric Hematology and Oncology, Manton Center for Orphan Disease Research, Howard Hughes Medical Institute, Children's Hospital Boston and Dana Farber Cancer Institute, Boston, MA, USA; [2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA; [3]Harvard Stem Cell Institute, Cambridge, MA, USA

Human diseases such as heart failure, diabetes, neurodegenerative disorders, and many others result from the deficiency or dysfunction of critical cell types. Strategies for therapeutic tissue repair or regeneration require the *in vitro* manufacture of clinically relevant quantities of defined cell types. In addition to transplantation therapy, the generation of otherwise inaccessible cells also permits disease modeling, toxicology testing and drug discovery *in vitro*. In this review, we discuss current strategies to manipulate the identity of abundant and accessible cells by differentiation from an induced pluripotent state or direct conversion between differentiated states. We contrast these approaches with recent advances employing partial reprogramming to facilitate lineage switching, and discuss the mechanisms underlying the engineering of cell fate. Finally, we address the current limitations of the field and how the resulting cell types can be assessed to ensure the production of medically relevant populations.

Introduction

Cell differentiation has long been thought to be a unidirectional process. The early development of all mammals involves a sequence of cell fate decisions along an irreversible pathway of restricted potential and increasing specialization. Development begins with the totipotent zygote, which has the capacity to give rise to all the embryonic and extraembryonic cells of the organism. After only a few days of development, specialized cell types emerge: trophectoderm and primitive endoderm form extraembryonic tissues supporting development, while epiblast constitutes the stem cells which ultimately give rise to hundreds of different somatic cell types and the germ lineage [1]. Epiblast can be coaxed to generate embryonic stem cells (ESCs) which are pluripotent, meaning they can give rise to cells of all three embryonic germ layers: ectoderm, mesoderm and endoderm, but not extraembryonic tissues. It was long believed that cell fate

could only transition to progressively more differentiated states, with de-differentiation seen only in cases of tissue pathology (e.g., metaplasia or malignancy). In the past half-century, this view has been challenged.

Fundamental discoveries in reprogramming cell fate

The advent of nuclear transplantation was a seminal moment in the goal to manipulate cell fate. In 1952, Briggs and King successfully transplanted intact nuclei from the amphibian *Rana pipiens* into an enucleated oocyte [2], in a technique later termed 'Somatic Cell Nuclear Transfer' (SCNT), or cloning, to imply perfect copying. Swimming tadpoles could be cloned from the nuclei of blastocysts, but more differentiated cells from gastrulation stages onwards generated aberrant tadpoles at best. It was thus concluded that the late gastrula nucleus has an 'intrinsic restriction in potentiality for differentiation' [3]. John Gurdon challenged these findings using an alternative amphibian model, *Xenopus laevis*, and demonstrated that the nucleus of a tadpole intestinal cell could give rise to a mature, fertile animal [4, 5]. Further work established that nuclei from terminally differenti-

Correspondence: George Q Daley
E-mail: george.daley@childrens.harvard.edu

ated cells could generate larvae, but not adults, following nuclear transfer (reviewed in [6]). These fundamental experiments demonstrated that the differentiated cell state is not a result of irreversible changes at the genomic level. In fact, the nuclei of somatic cells retain the capacity to orchestrate development into a fully functional organism.

Three decades passed before the success of cloning by SCNT could be recapitulated in a mammal. The arrival of Dolly the sheep, cloned by transfer of a mammary epithelial cell into an enucleated oocyte reignited interest in the field [7]. The first cloned mouse, 'Cumulina', followed a year later by transfer of a cumulus cell nucleus into an enucleated oocyte, although no other cell types tested in that study could generate mice [8]. The efficiency of successful cloning reported up to that point was 1%-2% (reviewed in [9]), which prompted speculation that the differentiated cell populations used to clone the animals may have been contaminated with a small proportion of stem cells, known to be more amenable to reprogramming [10]. To address this, nuclei were obtained from terminally differentiated cells that were genetically marked—adult B and T cells which have undergone genomic rearrangements at the immunoglobulin and T cell receptor loci—therefore providing proof of their maturity. Mice could not be immediately generated from these nuclei, thus an alternate strategy was employed. Blastocysts were first generated by nuclear transfer, followed by ESC derivation. Cloned mice were then generated by injecting the ESCs into tetraploid embryos, so that all tissues in the resulting mice were derived from the cloned ESCs, and indeed, all tissues displayed immunoglobulin or T cell receptor rearrangements [11]. SCNT into human oocytes had been unsuccessful until recently, when somatic nuclei were transferred into human oocytes in which the nucleus was left intact. These experiments generated triploid blastocysts from which stable triploid ESC lines could be derived. While they are not normal diploid hESCs, the triploid lines behave like typical pluripotent hESCs and thus establish that human oocytes can mediate reprogramming of somatic nuclei [12].

Directly switching cell fate without returning to a totipotent or pluripotent state

Experiments with nuclear transplantation revealed the capacity for a differentiated nucleus to be reprogrammed to a totipotent state, from which point it can develop into any cell type of the adult organism. This prompted researchers to investigate the potential for one mature somatic cell type to be directly converted into another mature somatic cell of alternate fate, bypassing an intermediate totipotent or pluripotent state.

To explore this potential, scientists fused two cells from different origins into heterokaryons, a kind of cell hybrid in which the nuclei remain distinct [13]. Experiments in which human amniocytes were fused to mouse muscle cells showed activation of human muscle-specific genes in the heterokaryons within 24 hours, demonstrating that gene expression repressed in differentiated cells could be reactivated [14]. Such fused cells have been extremely useful for our understanding of the mechanism of reprogramming. For example, these experiments have shown that nuclei of more specialized cells are generally more resistant to reprogramming, relative to less differentiated cells [15]. The same correlation between ease of reprogramming and differentiation status has been noted in SCNT experiments [10].

The pivotal experiments described thus far demonstrate that oocytes and somatic cells have the powerful capacity to direct cell fate through trans-acting reprogramming factors that regulate the epigenome of the cell. A critical contribution to the field was made by efforts to isolate these specific factors. A forerunner of this endeavor is represented by the work of Taylor and Jones, who treated immortalized fibroblasts with 5-azacytidine (5-Aza), an inhibitor of DNA methylation, and observed spontaneous differentiation into adipocytes and chondrocytes. This suggested that DNA methylation restricted gene expression of alternate lineages [16], and compelled Davis, Weintraub and Lassar to seek a gene responsible for the muscle fate switch in 5-Aza-treated fibroblasts. This led to the discovery that the transcription factor, MyoD could convert fibroblasts into contracting myocytes [17]. Further experiments to convert pigment, nerve, fat and liver cell lines produced cells that expressed muscle markers, but were aberrant due to their maintenance of starting cell type identity [18].

Transcription factor mediated reprogramming

The above knowledge and approaches were combined to explore whether reprogramming by nuclear transfer was mediated exclusively by trans-acting factors in the oocyte, or whether specific factors could reprogram cells in an ooplasm-independent manner. In support of the latter, fusion of ESCs with fibroblasts could reprogram the somatic cells to a pluripotent state, demonstrated by their capacity to differentiate into all three germ layers [19, 20]. These studies paved the way for the identification of specific reprogramming factors, akin to the identification of MyoD. The most successful attempt, and one of the most influential experiments in biology of the last decade, was performed by Takahashi and Yamanaka

who effected somatic cell reprogramming by combined expression of 24 ESC-specific genes in fibroblasts [21]. Then, through a process of elimination, they demonstrated that four transcription factors, Oct4, Sox2, Klf4 and c-Myc (OSKM) were sufficient to reprogram somatic cells to a pluripotent state, measured by the capacity of the cells to differentiate into all three germ layers, and contribute to embryonic development. The generation of such induced pluripotent stem cells (iPSCs) was a major breakthrough, demonstrating that just four factors can erase the epigenetic marks characteristic of a differentiated cell, and reset the cell to a pluripotent state. The field of induced pluripotency has since grown exponentially, and is a subject covered in detail by many excellent reviews [22-24].

The monumental demonstration of somatic cell reprogramming has prompted a resurgence of interest in direct conversion, a process by which cell fate is converted between two mature states by specific factors, without reverting to pluripotency [25-29] (Figure 1). Recently, novel strategies are emerging which combine partial reprogramming to pluripotency with direct conversion [30]. For the remainder of this review, we will make a comparison of the current progress in these areas, potential underlying molecular mechanisms, and possible limitations associated with each approach.

Directed differentiation

Self-renewal and pluripotency are the hallmarks of ESCs and iPSCs. The capacity for pluripotent cells to differentiate into the many cell types of the adult organism, coupled with their ability to be cultured and expanded *in vitro* [31-33] offers powerful new strategies for modeling human disease and developing personalized regenerative cell therapies. Derivation of iPSCs from adult cells also circumvents the ethical debate concerning the derivation of ESCs from human embryos. Harnessing this potential of pluripotent cells essentially relies on recapitulating development *in vitro* towards the desired *in vivo* cell type, termed 'directed differentiation'. In this regard, stem cell biologists have gleaned many cues from developmental biology.

The first step in engineering pluripotent cells towards the desired cell type is to guide their differentiation into the appropriate germ layer: ectoderm, mesoderm or endoderm. This is often achieved by adding specific embryonic morphogens or growth factors to the culture medium, such as Activin, Bone Morphogenetic Proteins (BMPs), WNTs (Int1, mammalian homologue of *Drosophila wingless*) and Fibroblast Growth Factors (FGFs) [25]. Further differentiation towards the desired end
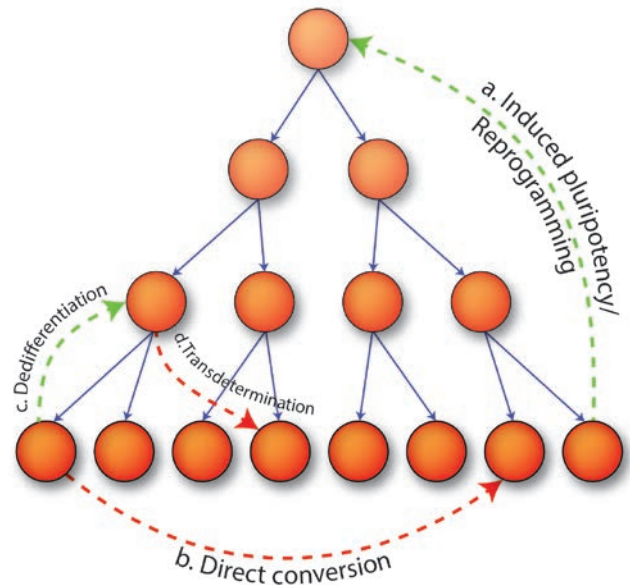

Figure 1 Routes of cell fate conversion. **(a)** Reversion to a totipotent state can be achieved by somatic cell nuclear transfer, whereas reversion to a pluripotent state is accomplished by iPS cell reprogramming, or cell fusion with ESCs. **(b)** Direct conversion of fate between two differentiated states without de-differentiation. Examples of direct conversion include fibroblast to neuron, hepatocyte or cardiomyocyte lineages. **(c)** De-differentiation to a progenitor state, as exemplified by loss of *Pax5* expression in mature B cells. **(d)** Transdetermination of an adult stem cell from its normal lineage into a closely related lineage. Each sphere represents a distinct stage of cell differentiation, with the most brightly colored spheres corresponding to mature cells. Directed differentiation directs cells down this natural cascade of fate specification.

point is achieved by additional growth factors, or small molecules acting on specific signaling pathways. Coculture systems have also been widely employed in an attempt to recapitulate the *in vivo* niche of the desired cell target. Numerous cell types have been produced through directed differentiation in normal and disease-specific contexts, as covered in many excellent reviews [34-37]. To illustrate such a directed differentiation strategy, two approaches are currently commonly employed to generate cardiomyocytes from pluripotent cells. In the first methodology, iPSCs are differentiated as embryoid bodies to promote initial differentiation into mesoderm, followed by treatment with a specific sequence of growth factors to guide the cells towards a cardiac fate [38]. Alternatively iPSCs can be cultured as a monolayer followed by sequential treatment with Activin A and BMP4 growth factors [39]. Typically though, these methods can be technically demanding, time consuming, and inefficient, which has fuelled investigation into alternative

strategies.

One of the major limitations of directed differentiation is the length of time it takes to first reprogram somatic cells to pluripotency and then subsequently direct them to the desired fate. Since these protocols constitute several stages, the efficiency with which the final cell type is generated can be low. This inefficiency is compounded by the fact that differentiation of iPSCs can vary widely among lines [40]. Moreover, cells within the same line also possess different differentiation propensities [41, 42]. Another major limitation is the nature of cells produced by directed differentiation: they are typically immature cells corresponding to embryonic stages of development, rather than fully mature adult cells [43-47]. Once transplanted *in vivo*, directly differentiated cells have been reported to mature; this has been utilized to advantage where pancreatic endoderm derived from hESCs efficiently generated glucose-responsive endocrine cells after a period of three months following transplantation into mice [48]. While this strategy may be appropriate for the end-goal of *in vivo* transplantation, for disease modeling and drug toxicology testing, it is essential to recapitulate the *in vivo* target as closely as possible. Finally, challenges exist to fully purify differentiated cells from pluripotent cells which have the potential to form teratomas [49], although the technology is moving away from the use of oncogenes and viral integration in an effort to address this. Taken together, these limitations have encouraged alternate means of fate conversion to be pursued.

Circumventing pluripotency by direct fate conversion

The early MyoD work [17], which established that cell fate can be converted without reversion to a pluripotent state, together with Takahashi and Yamanaka's demonstration [21] that fate can be reprogrammed with a combination of transcription factors, suggested that abundant and accessible cells such as fibroblasts might be used for conversion to any clinically relevant cell type. A major rationale behind this was that directly converting between somatic cell types, especially closely related cells, might involve less epigenetic remodeling, be more efficient, and produce mature cells [24].

Conversions in differentiated blood lineages have been informative with regards to the mechanism of direct conversion, as hematopoiesis is relatively well-defined [50]. In early work, ectopic expression of the erythroid-megakaryocyte associated transcription factor, GATA1, was shown to induce erythroid-megakaryocyte gene expression in monocytes (precursors to macrophages) [51]. Remarkably, expression of this single transcription

factor, GATA1, resulted in downregulation of monocytic markers [52, 53]. These experiments were originally performed in cultured cell lines, but were later shown to also apply to primary cells [54]. Conversely, introduction of the transcription factor and regulator of myeloid and B cell development, PU.1, into transformed multipotent hematopoietic progenitors repressed *GATA1* expression, leading to the upregulation of myeloid markers [55]. These experiments demonstrated the lineage-instructive role of transcription factors and underpin the ideology that transcription factor-mediated cell fate conversions mimic physiological cell fate transitions [27].

These early studies formed the groundwork to attempt direct conversion between mature hematopoietic lineages. Expression of the granulocyte/macrophage-specific transcription factor C/EBPα converted around 35% of primary mature B cells into macrophages, whereas 100% of less mature primary pro- and pre-B cells could be converted via this route [56]. Again, expression of this single transcription factor resulted in down-regulation of initial cell gene expression and up-regulation of target macrophage gene expression. Functionally, these induced macrophages demonstrated Fcγ receptor-dependent and -independent phagocytosis [56, 57]. Conversion efficiencies increased with the co-expression of C/EBPα and PU.1, where these factors in combination could convert more distant mesodermal cells, fibroblasts, but the resulting cells were only partially functional, and continued expression of the transgenes was required to maintain macrophage fate [58]. Taken together with the MyoD work, these studies suggest that expression of transcription factors represents a powerful methodology to produce stable fate changes in somatic cells.

Clinically relevant direct conversions

Many direct conversions between somatic cells have been reported, some of which have the potential to replenish defective or diminished cell types for therapy (Table 1). Examples include:

Differentiated exocrine pancreatic cells to β-cells

Pancreatic β-cells are a key target for the treatment of diabetes, due to their critical role in insulin storage and release. In a seminal paper, Zhou *et al*. screened expression of 1 100 transcription factors in mouse pancreatic tissue. 20 genes were found to be expressed in mature β-cells and their precursors: 9 of these genes produced phenotypes when mutated. Introduction of all 9 genes by adenovirus to the pancreas of immune-deficient mice resulted in an increase in the numbers of β-cells. This initial cocktail of genes was narrowed down to three fac-

Table 1 Compilation of direct conversions reported to date

Starting cells	Target cell type	Factors	References
Mesoderm			
Mouse Fibroblasts	Myocytes	MyoD	Davis *et al*., 1987 [17]
	Macrophages	CEBP/α/β, PU.1	Feng *et al*., 2008 [58]
	Cardiomyocytes	Gata4, Mef2c, Tbx5	Ieda *et al*., 2010 [65]
	Glutamatergic neurons	Ascl1, Brn2, Myt1l	Vierbuchen *et al*., 2010 [68]
	Dopaminergic neurons	Ascl1, Lmx1a, Nurr1	Caiazzo *et al*., 2011 [76]
	Motor Neurons	Brn2, Ascl1, Myt1l, Lhx3, Hb9, Isl1, Ngn2	Son *et al*., 2011 [82]
	Neuronal Precursor cells	Brn2, Sox2, Foxg1	Lujan *et al*., 2012 [77]
	Hepatocytes	Gata4, Hnf1α, Foxa3 P19^{ARF} knockdown	Huang *et al*., 2011 [79]
	Hepatocytes	Hnf4α, Foxa1/2/3	Sekiya and Suzuki, 2011 [80]
Human Fibroblasts	Glutamatergic neurons	Ascl1, Brn2, Myt1l, NeuroD1	Pang *et al*., 2011 [73]
	Glutamatergic neurons GABAergic neurons	Ascl1, Myt1l, NeuroD2, miR-9/9, miR-124	Yoo *et al*., 2011 [74]
	Glutamatergic neurons	Brn2, Myt1l, miR-124	Ambasudhan *et al*., 2011 [83]
	Neurons	Ascl1, Brn2, Myt1l, Zic1, Olig2	Qiang *et al*., 2011 [78]
	Dopaminergic neurons	Ascl1, Brn2, Myt1l, Lmx1a, Foxa2	Pfisterer *et al*., 2011 [75]
	Dopaminergic neurons	Ascl1, Lmx1a, Nurr1	Caiazzo *et al*., 2011 [76]
	Motor Neurons	Brn2, Ascl1, Myt1l, Lhx3, Hb9, Isl1, Ngn2	Son *et al*., 2011 [82]
	Hematopoietic progenitors	Oct4	Szabo *et al*., 2010 [84]
Mouse B Cells	Macrophages	CEBP/α/β	Xie *et al*., 2004 [56]
	T Cells	Pax5 knockout	Cobaleda *et al*., 2007 [85]
Mouse pre-T cells	Macrophages	CEBP/α/β	Laiosa *et al*,. 2006 [86]
	Dendritic cells	PU.1	Laiosa *et al*,. 2006 [86]
Ectoderm			
Mouse cortical astrocytes	Glutamatergic neurons	Ngn2	Heinrich *et al*,. 2010 [87]
	GABAergic neurons	Dlx2	Berninger *et al*., 2007 [88]
	GABAergic neurons	Ascl1, Dlx2	Heinrich *et al*,. 2010 [87]
Endoderm			
Mouse hepatocytes	Neurons	Ascl1, Brn2, Myt1l	Marro *et al*., 2011 [89]
Mouse and human hepatocytes	Pancreatic islet cells	See references	Ferber *et al*., 2000 [90], Kaneto *et al*., 2005 [91], Miyatsuka *et al*., 2003 [92] Sapir *et al*., 2005 [93] Yechoor *et al*., 2009 [63] Kojima *et al*., 2003 [94]
Mouse pancreatic exocrine cells	β-cells	Ngn3, Pdx1, MafA	Zhou *et al*., 2008 [59]

tors: Pdx1, Ngn3, and MafA [59]. Pdx1 plays an early role in pancreas development [60] and a later function in mature β-cell glucose homeostasis [61]. Ngn3 is a pro-neural transcription factor that is known to regulate development of the pancreatic endocrine lineages [62],

and its expression in the adult liver can convert hepatic progenitor cells to pancreatic islet tissue [63]. MafA is a transcription factor whose loss results in impaired glucose-stimulated insulin secretion, although β-cell specification is not abrogated, suggesting a late-stage function

[64]. Transient expression of these three transcription factors stably converted 20% of pancreatic acinar cells to β-cells, but had no effect when introduced into fibroblasts. The absence of *Sox9* and *Hnf6* expression indicates that cell fate was directly converted without transitioning through a progenitor state. The functionality of the converted cells was demonstrated using a type I diabetes mouse model where the *in vivo* direct conversion could alleviate insulin deficiency-induced hyperglycemia [59]. The success of this conversion hinged on the fact that it was performed *in vivo* as the induced β-cells could immediately reside in the native environment to support their survival and maturation.

Cardiomyocytes

Currently, the only treatment available for end-stage heart failure is whole-organ transplant, which is restricted by the availability of donor organs and the likelihood of immune rejection. With the goal of producing cells for replacement therapy, mouse cardiac and dermal fibroblasts have been converted to cardiomyocyte-like cells. Candidate factors were selected from a suite of genes expressed in cardiomyocytes and associated with clear developmental cardiac defects in mutants. From this initial pool, Gata4, Mef2c and Tbx5 were found to induce the expression of cardiac markers in fibroblasts without transitioning through a cardiac-progenitor state [65]. These transcription factors are known to form a core transcriptional network regulating cardiac development [66, 67]. Functionally, induced cardiomyocytes (iCMs) were electrophysiologically similar to ventricular cardiomyocytes, and a small population could simultaneously contract. Global gene expression analyses indicated that iCMs were similar, but not identical to neonatal cardiomyocytes [65]. This suggests that functional differences between cardiomyocytes and iCMs may exist, and that a period of maturation *in vivo* may be required.

Neurons

Due to their potential therapeutic value, direct conversion to neurons has been studied extensively (Table 1). Vierbuchen *et al*. successfully converted fibroblasts to neurons, representing a longer 'leap' in cell fate from mesoderm to ectoderm. From a pool of 18 selected candidate genes *Ascl1* expression induced neural characteristics in fibroblasts [68]. Ascl1 functions in mammals to regulate multi-potential stem cell differentiation in the central and peripheral nervous systems [69, 70]. Through systematic addition of the remaining transcription factors, Ascl1, Brn2 and Myt1l were found to directly convert embryonic and postnatal fibroblasts into neurons [68]. *Brn2* is expressed in cortical progenitors which

give rise to glutamatergic neurons [71] and is associated with cortical progenitor proliferation [72], and thus may explain why the induced neurons exhibit a glutamatergic phenotype. Functionally, the induced neurons could fire repeated action potentials and form synapses *in vitro*. This method has now been extended to the conversion of human fibroblasts to neurons [73-75]. Other milestones include the conversion of mouse and human fibroblasts to dopaminergic neurons [75, 76], the conversion of mouse fibroblasts to tripotent neural precursors [77] and derivation of neurons from fibroblasts of Parkinson's and Alzheimer's disease patients [76, 78].

Hepatocytes

Hepatocytes are valuable in terms of transplantation and drug toxicology testing, but a major limitation is that primary hepatocytes cannot be expanded *in vitro*. Recently, two independent studies [79, 80] focused on a suite of transcription factors known to be involved in liver development and hepatocyte function [81]. From these gene sets, combinations were identified by screening for *de novo* hepatic gene expression in transduced fibroblasts: Hnf4α and Foxa1/2/3 [80]; Gata4, Hnf1α, Foxa3 and inactivation of p19^{Arf} [79]. Although both groups claimed that progenitor marker expression was not detected in induced hepatocytes (iHeps), the resulting cells were immortalized and appeared to be only partially differentiated toward a liver fate. Functionally, iHeps were capable of engrafting fumarylacetoacetate hydrolase-deficient (*Fah*$^{-/-}$) mice, a tyrosinemia model subject to liver failure in the absence of drug treatment. Unlike primary hepatocytes though, iHeps were unable to fully repopulate *Fah*$^{-/-}$ mice and liver function did not appear to be normal, resulting in reduced survival rates. Interestingly, iHeps 'matured' after transplantation and ceased to express progenitor markers such as *Afp* [80]. In both studies, cell fate was stable following removal of the exogenous conversion factors [79, 80].

Mechanism of transcription factor mediated direct conversion

From the direct conversions described so far, transcription factors are predominantly responsible for driving fate change. Typically, the up-regulation of target gene expression is rapid, within hours to several days [65, 68]. Fate changes are direct, avoiding transition through a progenitor state [59, 65], and cell identity is stable after removal of exogenous factors [59, 79, 80]. Finally, fate conversion is achieved in the absence of cell division [59, 68, 87], in contrast to the induction of pluripotency which requires cell proliferation [95].

In the nucleus, the majority of DNA is packed into nucleosomes, occluded by higher order chromatin structure and repressors. Cell proliferation may facilitate reprogramming by allowing transcription factor access to otherwise occluded cis-regulatory regions through nucleosome displacement during DNA replication [28, 96]. Several models have been proposed to account for the access of transcription factors to their relevant binding sites to effect genome-wide transcriptional and epigenetic changes in the absence of cell division. One particularly favored model is the 'pioneer' transcription factor model. Pioneer factors can access their target sites in repressed regions of the genome where other factors cannot [97]. This is supported by the recurrence of such pioneer factors in the above conversion approaches: Gata4 [65, 79] and FoxA [75, 79, 80]. The FoxA factors for example can enable binding of other transcription factors that cannot bind in isolation [98]. This access is provided by local chromatin opening, nucleosome repositioning, and recruitment of chromatin modifiers and co-regulators [97]. This initiation is then followed by feed-forward induction of additional transcription factors to execute the differentiation process [99].

A recent elegant study has demonstrated how, genome-wide, a high percentage of Polycomb targets are associated with putative enhancers in permissive states, proposing a mechanism for the initiation of fate conversion and reprogramming [100]. In differentiated cells, genes can be expressed in a tissue-specific manner or repressed by Polycomb-repressive complex (PRC) and the associated H3K27me3 mark [101]. In ESCs, PRC targets are usually repressed but poised for activation, possessing both active (H3K4me3) and repressive (H3K27me3) marks [102, 103]. Taberlay *et al*. employed cell lines not normally expressing *MyoD1*, and found that a minimal *MyoD1* enhancer element existed in a 'permissive' state for MYOD1 binding (nucleosome-depleted, H3K4me1-enriched, not bound by PRC components or H3K27me3 enriched, flanked by H2AZ-containing nucleosomes), even in the presence of a repressive promoter state. This so-called 'multivalent' epigenetic state at the *MYOD1* locus permitted ectopic MYOD1 binding to the permissive enhancer within 24 hours of its expression, resulting in nucleosome displacement at the promoter and the emergence of H3K4me3 enrichment by 48 hours. Interestingly, although ths is insufficient to induce transcription from the *MYOD1* locus, addition of conditioned medium from a *MYOD1*-expressing rhabdomyosarcoma cell line resulted in *MYOD1* transcription, demonstrating the importance of contextualizing signals. In contrast, in a cell line where both the enhancer and promoter were nucleosome bound, MYOD1 binding was not observed. This

permissive enhancer state is common throughout the genome, suggesting one mechanism by which cells retain epigenetic plasticity [100]. This opens up the possibility of utilizing such datasets to predict the best starting cell populations that will respond to specific combinations of transcription factors.

Direct conversion efficiencies

In reprogramming to pluripotency, the differentiation status of the starting cell has been shown to influence reprogramming efficiency [24]. In agreement with nuclear transfer and cell fusion studies discussed earlier, where less differentiated cells can be reprogrammed with greater ease, the same also applies to transcription factor mediated reprogramming. Mouse-derived neural stem cells generate iPSCs 50-fold more efficiently than fibroblasts [104]. Also, differentiated cells have demonstrated different reprogramming efficiency properties with keratinocytes reprogramming faster than fibroblasts [105], and with higher efficiency [106].

Considering the potential differences in reprogramming efficiencies depending on cell-type of origin, the same trend is likely to apply to direct conversion. For example, closely related cell types, which are more similar epigenetically, may convert more efficiently. Evidence supports this notion: MyoD can convert fibroblasts into contracting myocytes, both mesodermal tissues, with 25-30% efficiency [17]. In contrast, when MyoD was expressed in cells from different germ layers (pigment cells, melanocytes or hepatocytes) reprogramming was incomplete, producing aberrant cells at low frequencies [18]. The notion is also supported by the high efficiency conversion of mature B-cells into macrophages after CEBP/α overexpression (35% efficiency) [56], which contrasts with the incomplete conversion of fibroblasts to macrophages with PU.1 and CEBP/α, which cannot be maintained after removal of exogenous factors [58]. Based on current experience, the feasibility of transcription factor-mediated direct conversion across germ layers seems limited, but warrants further investigation.

Epigenetic memory and evidence for incomplete conversion

Transcriptional remnants or residual chromatin features characteristic of the starting cell type that persist after reprogramming, potentially impacting target cell function, represent "epigenetic memory". Residual epigenetic memory was first documented in embryos cloned from *Sox2*-expressing neuroectodermal nuclei. 81% of these embryos aberrantly expressed *Sox2* in endoderm

[107], suggesting that an epigenetic mark of transcriptional activity was maintained through the reprogramming process. This concept was further supported by a report of residual epigenetic memory of MyoD in cloned embryos, correlating with K4-trimethylated H3.3 retention at the *MyoD* promoter [108]. Moreover, iPSCs sometimes exhibit residual epigenetic memory associated with the cell type of origin [109-111].

Residual epigenetic memory also complicates direct conversion; in the case of induced neural cells derived from hepatocytes, while most hepatic genes were down-regulated, some hepatocyte-specific expression persisted [89]. Additionally, following conversion of fibroblasts to macrophages, some fibroblast gene expression has been documented to persist. Moreover, the resulting cells are an unstable macrophage intermediate, de-differentiating after removal of the exogenous factors [58]. Finally, direct conversion of fibroblasts to cardiomyocytes produces cells that do not fully recapitulate the profile of neonatal cardiomyocytes, which could be accounted for by an epigenetic memory of initial host cell identity [65].

The reports of persisting gene expression raise the possibility that the desired identity of the target cell types may not be fully achieved by these approaches, in terms of both silencing host cell gene expression and establishing target cell gene regulatory networks (GRNs). Cell identity is determined by these genetically defined and epigenetically modulated GRNs, which describe the transcriptional relationships among genes [112]. A major issue in the analysis of the cell conversions reported so far is that often only a handful of cell markers are analyzed and the establishment of any alternative GRNs, for example those that would arise from the reuse of transcriptional programs are not assessed. This is a critical point as some of the pioneer transcription factors employed, such as Gata4 and FoxA factors can initiate different genetic programs throughout development [65, 75, 79, 80], which could explain why several factors are often required to elicit stable fate changes. We have recently developed a computational platform to help address these issues. By reconstructing GRNs of many cell types and tissues, we can identify regulatory nodes at which engineered cells are distinct from target cell types, thus allowing for a metric of engineered cell identity. From our analyses of the fibroblast to hepatocyte conversion driven by Hnf4α and Foxa1 [80], we find that fibroblast identity is not silenced, and target hepatocyte gene expression is not fully realized, relative to primary hepatocytes. These findings may explain why iHeps are not functionally equivalent to primary hepatocytes, a critical point to address as equivalence to *in vivo* hepatocytes is essential for *in vitro* drug toxicology studies.

Together, the many limitations noted suggest that the direct conversion strategy may not precisely recapitulate target cell identity, and may at best be limited to conversions between closely related cell types. There is evidence from cell-fusion studies to support this concern. Following fusion of myotubes and fibroblasts, some myogenic genes are activated more slowly, or not at all, in the fibroblast nuclei [113]. This form of epigenetic silencing is termed 'occlusion'. Occluded genes are silenced by cis-acting epigenetic modifications irrespective of the presence of trans-acting factors. It is worth noting that in heterokaryon experiments, the dose of trans-acting factors is much lower relative to forced transcription factor expression. Even so, this may explain the residual epigenetic memory of directly converted cells, which appears to be more prominent, relative to iPSCs. Direct conversion offers benefits of higher efficiencies, greater speed and potential safety due to the avoidance of pluripotent intermediates. Regardless, a major drawback is the limited expansion potential of the target cells, and the possibility that it is very difficult to fully recapitulate target cell types. Thus we must continue to consider the induction of pluripotency as a real alternative.

Cell fusion with ESCs reactivates pluripotency [19], which appears dominant over differentiation [20, 114], and these results suggest that a return to the pluripotent state is unique in that it releases most genes from occlusion. This is supported by evidence from the preimplantation embryo in which pluripotent cells undergo a wave of demethylation of the Polycomb mark H3K27me3 [115], an epigenetic mark associated with occlusion. Considering this, induction of pluripotency may represent a more powerful mechanism to erase the old epigenome and enable cell fate changes. Recently, interesting new strategies are emerging which may exploit this mechanism, in combination with direct conversion, potentially offering the benefits of both strategies.

An emerging alternative strategy: A bypass to cardiomyocytes

There has been consistent interest in deriving cardiomyocytes from a more abundant source, although challenges are faced with current approaches. Directed differentiation from iPSCs is time-consuming since somatic cells first have to be reprogrammed to pluripotency, and subsequently differentiated to cardiomyocytes. As discussed earlier, this is technically challenging and laborious; the process is generally inefficient and it is difficult to remove pluripotent cells that could give rise to teratomas. There has also been a high degree of variability in cardiac differentiation capacities reported for individual

pluripotent lines [38]. Considering these limitations, Ieda *et al.* generated cardiomyocytes from fibroblasts using a direct conversion strategy (see earlier [65]), although the global transcriptional analysis of these cells demonstrated that there are differences, compared to neonatal cardiomyocytes, raising the possibility that they are not true functional equivalents.

Recently, an alternative strategy for the conversion of fibroblasts to cardiomyocytes was reported [116]. In this study, four days of ectopic Oct4, Sox2, Klf4 and c-Myc (OSKM) expression in tail tip fibroblasts was employed as a 'shortcut' to mouse cardiogenesis. This bypassed a pluripotent intermediate while supporting a transient, plastic developmental state known to be established early in reprogramming [117]. This essentially functioned as a 'springboard' for subsequent conversion to cardiomyocytes under culture conditions supporting cardiac differentiation. To divert cells away from a pluripotent state, the culture was supplemented with an inhibitor of JAK/STAT signaling. To further argue against the acquisition of pluripotency which could be responsible for promoting cardiogenesis under these conditions, the presence of LIF delayed the onset of beating and decreased the number of beating colonies. Moreover, OSKM expression induction is limited to only 4 days, whereas it is known to take at least 12 days for iPSC colonies to emerge [118, 119]. This is also supported by the absence of Nanog expression during the conversion process. Interestingly, c-Myc was dispensable for the conversion of less mature mouse embryonic fibroblasts to cardiomyocytes, in agreement with the data suggesting that less differentiated cells are easier to reprogram [111].

Is such a 'primed conversion' approach more efficient than conventional reprogramming methods and does it yield cell types that more faithfully resemble their *in vivo* counterparts? Ieda *et al.* directly converted fibroblasts into 'cardiomyocyte-like cells' by overexpression of the cardiogenic transcription factors *Gata4, Mef2c*, and *Tbx5* in cardiac and dermal fibroblasts [65]. The resulting cells expressed cardiac markers and spontaneously contracted. 10% of fibroblasts can be converted to cardiomyocytes (as measured by the expression of cardiac troponin T, a marker of mature cardiomyocytes) compared to 40% by the primed conversion approach [116]. Following direct conversion, no markers of cardiovascular progenitors (such as *Mesp1, Nkx2.5, Flk1 and Gata4*) were reported, in contrast to the primed conversion approach where these progenitor markers were detected. It must be noted though, that these "primed" cells may be endowed with short-term self-renewal capacity prior to their differentiation, thus accounting for the higher efficiency of conversion. It must also be considered that the different starting cell populations used in these two approaches may complicate direct comparison of the methods. In addition the cells subjected to primed conversion generate atrial cardiomyocytes (as shown by expression of the atrial form of myosin light chain isoform [116]). Ventricular cardiomyocytes would be more desirable as these are the cells in the heart that take the heavy work load. Although direct conversion generates cardiomyocytes with action potentials similar to adult ventricular cardiomyocytes, the transcriptional profiles are distinct, suggesting functional differences exist [65]. Taking cues from directed differentiation, it may in the future be possible to direct the primed cells towards a ventricular fate. It would also be interesting to further explore this strategy to see if the engineered cells could better recapitulate *in vivo* populations.

Dual roles of Oct4

The Efe *et al.* study suggests that a brief exposure to reprogramming factors converts cells from a stable differentiated state into a transient, epigenetically unstable state [116]. It has been suggested that such a partially reprogrammed stage represented in the first stages of this primed conversion, or indirect conversion [30], approach parallels the *in vivo* dedifferentiation observed in the regeneration process of some lower vertebrate species [120]. In the transition toward the pluripotent state, somatically acquired epigenetic marks are erased, facilitating transition to an alternate cell fate. Oct4 is possibly the most critical of the reprogramming factors in this respect [24]. OCT4 association with PRC targets in pluripotent cells is associated with gene inactivity and a bivalent epigenetic promoter signature, characterized by both H3K4me3 and H3K27me3 marks. Such a bivalent signature in ESCs poises developmentally critical genes for activation, while the genes remain transcriptionally inactive [103]. To investigate the consequence of ectopic Oct4 expression in differentiated cells, Taberlay *et al.* identified an OCT4 binding site in the *MYOD1* enhancer (see above and Figure 2A). Upon introduction of ectopic Oct4 into fibroblasts, OCT4 occupancy at the enhancer was detected within 24 hours. By 72 hours, OCT4 binding could be detected at the *MYOD1* promoter, which consequently reverted to a bivalent state normally associated with pluripotency [100]. Considering these data in the context of primed conversion, we speculate that reprogramming factors such as Oct4, in addition to modifying the host cell epigenome, poise key fate genes for activation. The kinetics of promoter bivalent state acquisition (72 hours) certainly lie within the time period required for OSKM action in the primed conversion to

cardiomyocytes (96 hours) [116]. Following subsequent exposure to cardiogenic growth factors rather than pluripotency-promoting conditions, cardiac-specific transcriptional programs can be activated. While it is tempting to surmise that the reprogramming factors, Oct4 in particular, are simply acting to create an epigenetically permissive state, more complex mechanisms may be operating.

Oct4 dictates three fate outcomes in mouse ESCs: 1) A basal level of Oct4 corresponds with self-renewal and pluripotency; 2) Repression of Oct4 leads to trophectoderm differentiation; and 3) Less than a two-fold increase in Oct4 promotes differentiation into primitive endoderm and mesoderm [121]. Furthermore, it has been demonstrated that Oct4 dosage can regulate specification

of ESCs toward cardiogenesis, and that Oct4 expression in the late blastocyst is required for heart development *in vivo* [122]. In hESCs, OCT4 also acts as a dose-dependent switch regulating the transition from pluripotency to induction of cardiogenesis, hinging on its interaction with the transcription factors SOX2 and SOX17. It has been suggested that the basal level of OCT4 in hESCs targets the *OCT4-SOX2* enhancer, thus maintaining the *OCT4/SOX2/NANOG* loop, the core of a transcriptional network promoting pluripotency and self-renewal [123]. High OCT4 or low SOX2 levels permit OCT4 binding to the *SOX17* promoter. This activates *SOX17* expression, driving cells towards the endo/mesendodermal fate which functions within the ESC colony to induce cardiac
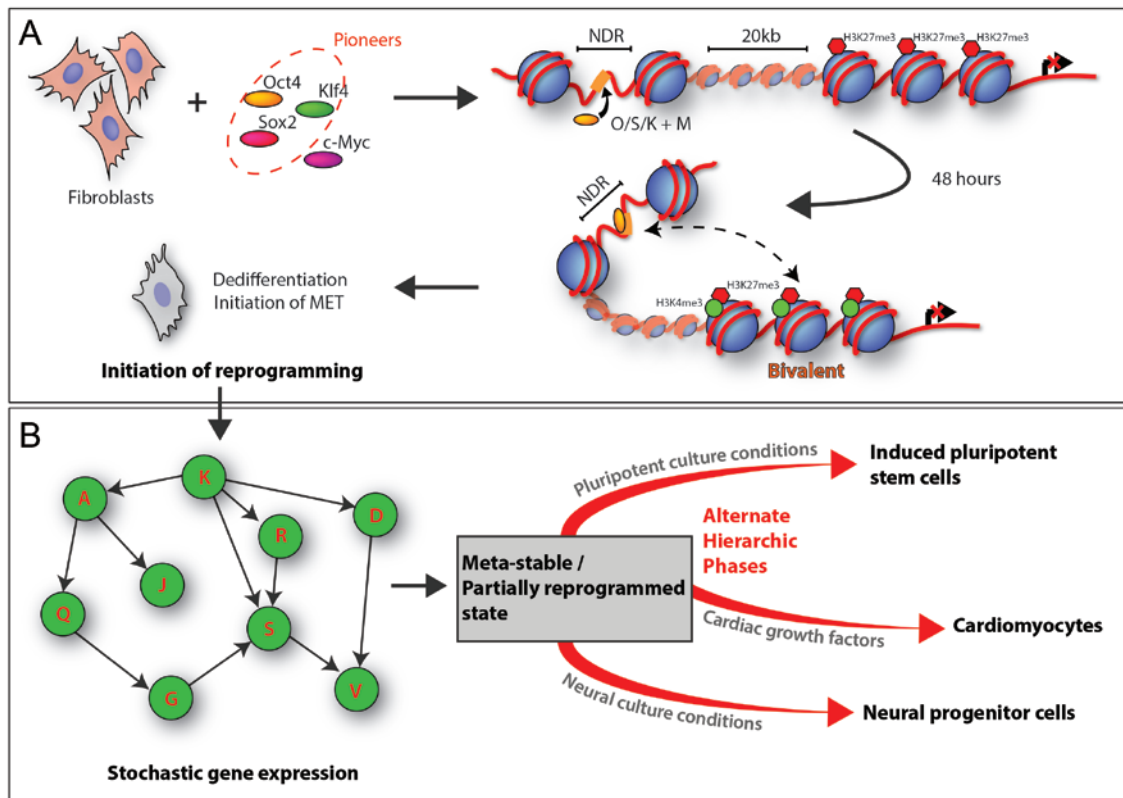

Figure 2 Proposed model for primed conversion. **(A)** Introduction of Oct4, Sox2, Klf4 and c-Myc into fibroblasts initiates reprogramming. OCT4, SOX2, KLF4 act as pioneer factors, facilitated by C-MYC [137] where OCT4 has been shown to bind to permissive enhancers to mediate reversion to a bivalent state at the promoter [100], poising key fate genes for activation, in addition to rapidly silencing the host epigenome. We speculate that this may help facilitate subsequent expression of occluded genes. **(B)** In these early stages of reprogramming, gene expression is stochastic [135] where reprogramming factors initiate a sequence of probabilistic events that lead to unpredictable and a small frequency of iPSC generation. Many key differentiation genes are known to be expressed in this partially reprogrammed state which may act as a platform for subsequent directed differentiation into discrete cell fates. For example, under conditions promoting reversion to pluripotency, a small fraction of the initial cell population may enter a hierarchic phase of ordered gene expression supporting iPSC formation. Under alternate culture conditions (e.g. inclusion of cardiac growth factors), partially reprogrammed cells may enter an alternate hierarchic phase cumulating in differentiation toward cardiomyocytes. Blue spheres: Nucleosomes. Red strand: DNA. NDR: Nucleosome Depleted Region.

fate (by secreting cardiogenic factors such as WNT3A and BMP2).

Efe *et al*. speculate that reprogramming factors, particularly Oct4, function to erase cell identity by epigenetic mechanisms, and do not directly activate lineage-specific genes [116]. They suggest that the reprogramming factors induce a 'developmentally more naïve, open-chromatin state marked by high epigeneticin instability'. This may certainly be true for the early stages of this primed conversion process, but it is possible that Oct4 levels in some colonies then push the cells towards cardiogenesis (as would be the case for ES cells), which is stimulated further by the culture conditions. This raises the possibility that the reprogramming process has the potential to be pushed in any direction. Interestingly, a reciprocal activity has been reported for Sox2, where a less than two-fold increase in its expression leads ESCs to predominantly differentiate into neuroectoderm, followed by mesoderm and trophectoderm, but not endoderm [124]. In a separate study, neuroectoderm predominated following Sox2 overexpression and culture under differentiation conditions [125]. Thus the stoichiometry of these reprogramming factors is clearly critical.

Adding further support to the concept of primed conversion, four-factor partial reprogramming has also been utilized to convert fibroblasts to neural progenitors. These neural progenitor cells can be expanded somewhat *in vitro* and give rise to several neuronal subtypes and glial cells, although oligodendrocyte differentiation may be limited [126]. More recently, constitutive Sox2, Klf4 and c-Myc expression with restricted Oct4 expression produced induced neural stem cells possessing tri-potentiality and seemingly unlimited expansion potential [127]. This approach potentially relies on the reuse of Sox2 and Klf4 in neural development. The reported conversion of mouse fibroblasts to hyaline cartilaginous tissue may also hinge on priming by c-Myc and Klf4 [128]. Furthermore, expression of OCT4 in combination with culture in hematopoietic cytokines (Stem Cell Factor and FMS-like tyrosine kinase 3 ligand) has been described to directly convert human fibroblasts to CD45[+]/CD34[+] hematopoietic progenitor colonies possessing erythroid and myeloid potential [84]. The authors proposed that OCT4 may directly interact with hematopoietic genes, based on its homology with OCT1/2, known lymphopoietic factors [129], but this seems unlikely as Oct4 is not required for hematopoiesis [130] and there is no evidence for ectopic hematopoiesis following forced Oct4 expression in adult mice [131]. One alternate possibility is that the formation of an endo/mesodermal intermediate could be promoted (akin to OCT4 overexpression in hESCs, above) which was not detected by the frequency of sampling in the study [84]. Another likely alternative is that Oct4 acting alone may destabilize the epigenome, leading to stochastic hematopoietic differentiation at a low frequency. It may help to understand the process of reprogramming to pluripotency and the formation of partially reprogrammed cells to propose a model for these fate transitions.

A model for 'primed conversion'

Reprogramming had previously been proposed to progress through three distinct phases: initiation, maturation and stabilization [132]. Microarray analysis has shown that mouse embryonic fibroblasts immediately de-differentiate when exposed to reprogramming factors [133]. In contrast to the relatively deterministic processes of nuclear transfer and cell fusion [23, 134], Hanna *et al*. proposed a stochastic model for pluripotency induction where reprogramming factors initiate a sequence of probabilistic events that lead to an unpredictable and small frequency of success in iPSC generation [95]. Technical limitations previously restricted studies to the clonal population level [117]. Buganim *et al*. recently studied the reprogramming process at the single cell level by profiling expression of 48 genes, including ESC chromatin modifiers, cell cycle regulators, signal transducers and pluripotency markers. To precisely track events, individual cells were clonally expanded and sister cells profiled [135].

In the earliest stages of reprogramming, no obvious sequential order of expression of the 48 genes tested was identified, supporting a stochastic mechanism for this initial phase. Unexpectedly, *Oct4* expression at this stage was not predictive for which cells would reach a pluripotent state. Rather, expression of *Esrrb*, *Utf1*, *Lin28*, and *Dppa2* appears to be a more reliable predictor for the ultimate acquisition of pluripotency. Surprisingly, expression of *Sox2* in the later stages could predict which cells would eventually become fully reprogrammed, from which point the process becomes deterministic. Impressively, the insight of this 'hierarchical' phase enabled reprogramming with *Esrrb*, *Utf1*, *Lin28*, and *Dppa2* in the absence of the four original factors [135].

Knowledge of molecular events in reprogramming can help to build a model of how primed conversion may be feasible. The initial phase of primed conversion requires a short exposure of differentiated cells to reprogramming factors [116]. During this phase, in addition to rapid de-differentiation [133], Oct4, for example, may act through binding to permissive enhancers to poise genes of alternate lineages for expression [100]. This first four days of the process would effectively represent the stochastic

phase [135] rendering the cells partially reprogrammed [117]. At this stage, a strong upregulation of lineage-specific genes from unrelated lineages has been reported [133]; Axon guidance: *Epha7* and *Ngef*; Epidermal proteins: *Krt14, Krt16, Ivl* and *Sprr1a*, and glomerular protein *Podxl*. This is reflective of the stem-cell independent roles of Sox2 and Klf4 in neural, epidermal and kidney development [136], and supports the notion that partially reprogrammed cells could be poised toward several differentiation states, depending on reprogramming factor stoichiometry.

We propose that in a partially reprogrammed, or 'primed' state, cell fate can be diverted toward alternate identities in a culture condition-dependent manner (Figure 2). Oct4 plays a key role in this process, supported by the fact that Oct4 is not a reliable indicator of pluripotency acquisition in the stochastic stages of reprogramming. From this meta-stable state, under conditions promoting stem cell formation, some cells enter the hierarchical phase towards pluripotency. Under modified conditions such as those used in the Efe *et al.* study to promote cardiac differentiation, some cells enter an alternate hierarchical phase cumulating with cardiomyocyte fate [116]. It is very likely also that the direction of fate change is influenced by the stoichiometry of the reprogramming factors themselves. This approach has the potential to activate expression of previously occluded genes, and this will form an essential avenue of future study. By undertaking a single cell profiling approach, the hierarchical phase of fate conversion whether it be to a pluripotent or fully differentiated state, could be dissected to understand the combinations of factors that may directly promote deterministic fate specification. This knowledge could be used to improve fate conversion efficiency and perhaps recapitulate the *in vivo* cell type more precisely. The process would have further potential to be tailored by initial selection of alternate reprogramming factors [24], depending on the nature of their reuse throughout development.

Future directions toward a blueprint of cell fate

As we have discussed in this review, there remain many limitations in our current attempts to engineer cell fate. Directed differentiation from a pluripotent state, while providing an abundant source of cells is time consuming, laborious, inefficient and potential safety concerns exist. Direct conversions between adult fates may be faster and more efficient, but has the disadvantage of a lack of scalability, which would be essential to offer any real therapeutic benefit. Moreover, there is evidence that fate transitions between converted somatic cell popula-

tions are not complete, with target cell identity not fully achieved. Primed conversion represents a middle ground between these two approaches. It remains to be proven though that the cell types produced by primed conversion are closer to their *in vivo* counterparts, which represents a critical future area of study. Notwithstanding the recent progress and explosion of attention being paid to cell fate conversions in biomedical research, more discriminating and comprehensive analyses of the molecular identity of target cell populations are needed, especially if we are to ensure safety and functionality in transplanted cell populations. Realizing the full potential of manipulating cell fates for biomedical applications will depend on refining our methods for cellular alchemy.

Note added in proof
In support of the proposed primed conversion model, Polo *et al.* [138] analyzed transcriptional and epigenetic changes in phenotypically defined iPSC intermediates. Transient changes in expression developmental regulators was observed, indicative of the primed, partially reprogrammed state. Furthermore, a primed conversion approach has recently been employed to convert human fibroblasts to angioblast-like progenitor cells [139].

References

1　Zernicka-Goetz M, Morris SA, Bruce AW. Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. *Nat Rev Genet* 2009; **10**:467-477.

2　Briggs R, King TJ. Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. *Proc Natl Aca Sci USA* 1952; **38**:455-463.

3　King TJ, Briggs R. Changes in the nuclei of differentiating gastrula cells, as demonstrated by nuclear transplantation. *Proc Natl Aca Sci USA* 1955; **41**:321-325.

4　Gurdon JB, Elsdale TR, Fischberg M. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* 1958; **182**:64-65.

5　Gurdon JB, Uehlinger V. "Fertile" intestine nuclei. *Nature* 1966; **210**:1240-1241.

6　Gurdon JB. From nuclear transfer to nuclear reprogramming: the reversal of cell differentiation. *Annu Rev Cell Dev Biol* 2006; **22**:1-22.

7　Campbell KH, McWhir J, Ritchie WA, Wilmut I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* 1996; **380**:64-66.

8　Wakayama T, Perry AC, Zuccotti M, Johnson KR, Yanagi-

machi R. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* 1998; **394**:369-374.

9   Gurdon JB, Byrne JA. The first half-century of nuclear transplantation. *Proc Natl Aca Sci USA* 2003; **100**:8048-8052.

10   Wakayama T, Rodriguez I, Perry AC, Yanagimachi R, Mombaerts P. Mice cloned from embryonic stem cells. *Proc Natl Aca Sci USA* 1999; **96**:14984-14989.

11   Hochedlinger K, Jaenisch R. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* 2002; **415**:1035-1038.

12   Noggle S, Fung HL, Gore A, *et al*. Human oocytes reprogram somatic cells to a pluripotent state. *Nature* 2011; **478**:70-75.

13   Blau HM, Blakely BT. Plasticity of cell fate: insights from heterokaryons. *Semin Cell Dev Biol* 1999; **10**:267-272.

14   Blau HM, Chiu CP, Webster C. Cytoplasmic activation of human nuclear genes in stable heterocaryons. *Cell* 1983; **32**:1171-1180.

15   Blau H, Pavlath GK, Hardeman EC, *et al*. Plasticity of the differentiated state. *Science* 1985; **230**:758-766.

16   Taylor SM, Jones PA. Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. *Cell* 1979; **17**:771-779.

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

18   Weintraub H, Tapscott SJ, Davis RL, *et al*. Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc Natl Aca Sci USA* 1989; **86**:5434-5438.

19   Cowan CA, Atienza J, Melton DA, Eggan K. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* 2005; **309**:1369-1373.

20   Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T. Nuclear reprogramming of somatic cells by *in vitro* hybridization with ES cells. *Curr Biol* 2001; **11**:1553-1558.

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

22   Robinton DA, Daley GQ. The promise of induced pluripotent stem cells in research and therapy. *Nature* 2012; **481**:295-305.

23   Hanna JH, Saha K, Jaenisch R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell* 2010; **143**:508-525.

24   Hochedlinger K, Plath K. Epigenetic reprogramming and induced pluripotency. *Development* 2009; **136**:509-523.

25   Cohen DE, Melton D. Turning straw into gold: directing cell fate for regenerative medicine. *Nat Rev Genet* 2011; **12**:243-252.

26   Graf T. Historical origins of transdifferentiation and reprogramming. *Cell Stem Cell* 2011; **9**:504-516.

27   Graf T, Enver T. Forcing cells to change lineages. *Nature* 2009; **462**:587-594.

28   Vierbuchen T, Wernig M. Molecular roadblocks for cellular reprogramming. *Mol Cell* 2012; **47**:827-838.

29   Vierbuchen T, Wernig M. Direct lineage conversions: unnatural but useful? *Nat Biotechnol* 2011; **29**:892-907.

30   Sancho-martinez I, Baek SH, Izpisua Belmonte JC. Lineage

conversion methodologies meet the reprogramming toolbox. *Nat Cell Biol* 2012; **14**:892-899.

31   Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 1981; **292**:154-156.

32   Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Aca Sci USA* 1981; **78**:7634-7638.

33   Thomson JA, Itskovitz-Eldor J, Shapiro SS, *et al*. Embryonic stem cell lines derived from human blastocysts. *Science* 1998; **282**:1145-1147.

34   Cherry ABC, Daley GQ. Reprogramming cellular identity for regenerative medicine. *Cell* 2012; **148**:1110-1122.

35   Teo AKK, Vallier L. Emerging use of stem cells in regenerative medicine. *Biochem J* 2010; **428**:11-23.

36   Wu SM, Hochedlinger K. Harnessing the potential of induced pluripotent stem cells for regenerative medicine. *Nat Cell Biol* 2011; **13**:497-505.

37   Nelson TJ, Martinez-Fernandez A, Terzic A. Induced pluripotent stem cells: developmental biology to regenerative medicine. *Nat Rev Cardiol* 2010; **7**:700-710.

38   Kattman SJ, Witty AD, Gagliardi M, *et al*. Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell* Stem *Cell* 2011; **8**:228-240.

39   Paige SL, Osugi T, Afanasiev OK, Pabon L, Reinecke H, Murry CE. Endogenous Wnt/beta-catenin signaling is required for cardiac differentiation in human embryonic stem cells. *PloS One* 2010; **5**:e11134.

40   Osafune K, Caron L, Borowiak M, *et al*. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat Biotechnol* 2008; **26**:313-315.

41   Chambers I, Silva J, Colby D, *et al*. Nanog safeguards pluripotency and mediates germline development. *Nature* 2007; **450**:1230-1234.

42   Fischer Y, Ganic E, Ameri J, Xian X, Johannesson M, Semb H. NANOG reporter cell lines generated by gene targeting in human embryonic stem cells. *PloS One* 2010; **5**:e12533.

43   Oldershaw RA, Baxter MA, Lowe ET, *et al*. Directed differentiation of human embryonic stem cells toward chondrocytes. *Nat Biotechnol* 2010; **28**:1187-1194.

44   Oshima K, Shin K, Diensthuber M, Peng AW, Ricci AJ, Heller S. Mechanosensitive hair cell-like cells from embryonic and induced pluripotent stem cells. *Cell* 2010; **141**:704-716.

45   D'Amour KA, Bang AG, Eliazer S, *et al*. Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat Biotechnol* 2006; 24:1392-1401.

46   Touboul T, Hannan NR, Corbineau S, *et al*. Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. *Hepatology* 2010; **51**:1754-1765.

47   Rashid ST, Corbineau S, Hannan N, *et al*. Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. *J Clin Invest* 2010; **120**:3127-3136.

48   Kroon E, Martinson LA, Kadoya K, *et al*. Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. *Nat Biotechnol* 2008; **26**:443-452.

49 Miura K, Okada Y, Aoi T, *et al*. Variation in the safety of induced pluripotent stem cell lines. *Nat Biotechnol* 2009; **27**:743-745.

50 Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* 2008; **132**:631-644.

51 Pevny L, Simon MC, Robertson E, *et al*. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* 1991; **349**:257-260.

52 Visvader JE, Elefanty AG, Strasser A, Adams JM. GATA-1 but not SCL induces megakaryocytic differentiation in an early myeloid line. *EMBO J* 1992; **11**:4557-4564.

53 Kulessa H, Frampton J, Graf T. GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblasts, and erythroblasts. *Genes Dev* 1995; **9**:1250-1262.

54 Heyworth C, Pearson S, May G, Enver T. Transcription factor-mediated lineage switching reveals plasticity in primary committed progenitor cells. *EMBO J* 2002; **21**:3770-3781.

55 Nerlov C, Graf T PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. *Genes Dev* 1998; **12**:2403-2412.

56 Xie H, Ye M, Feng R, Graf T. Stepwise reprogramming of B cells into macrophages. *Cell* 2004; **117**:663-676.

57 Di Tullio A, Vu Manh TP, Schubert A, Castellano G, Månsson R, Graf T. CCAAT/enhancer binding protein alpha (C/EBP(alpha))-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation. *Proc Natl Aca Sci USA* 2011; **108**:17016-17021.

58 Feng R, Desbordes SC, Xie H, *et al*. PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells. *Proc Natl Aca Sci USA* 2008; **105**:6057-6062.

59 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.

60 Jonsson J, Carlsson L, Edlund T, Edlund H. Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature* 1994; **371**:606-609.

61 Ahlgren U, Jonsson J, Jonsson L, Simu K, Edlund H. beta-cell-specific inactivation of the mouse Ipf1/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes. *Genes Dev* 1998; **12**:1763-1768.

62 Gu G, Dubauskaite J, Melton Da. Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development* 2002; **129**:2447-2457.

63 Yechoor V, Liu V, Espiritu C, *et al*. Neurogenin3 is sufficient for transdetermination of hepatic progenitor cells into neo-islets *in vivo* but not transdifferentiation of hepatocytes. *Dev Cell* 2009; **16**:358-373.

64 Zhang C, Moriguchi T, Kajihara M, *et al*. MafA is a key regulator of glucose-stimulated insulin secretion. *Mol Cell Biol* 2005; **25**:4969-4976.

65 Ieda M, Fu JD, Delgado-Olguin P, *et al*. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* 2010; **142**:375-386.

66 Olson EN. Gene regulatory networks in the evolution and development of the heart. *Science* 2006; **313**:1922-1927.

67 Srivastava D. Making or breaking the heart: from lineage determination to morphogenesis. *Cell* 2006; **126**:1037-1048.

68 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.

69 Nieto M, Schuurmans C, Britz O, Guillemot F. Neural bHLH genes control the neuronal versus glial fate decision in cortical progenitors. *Neuron* 2001; **29**:401-413.

70 Lo L, Sommer L, Anderson DJ. MASH1 maintains competence for BMP2-induced neuronal differentiation in post-migratory neural crest cells. *Curr Biol* 1997; **7**:440-450.

71 Leone DP, Srinivasan K, Chen B, Alcamo E, McConnell SK. The determination of projection neuron identity in the developing cerebral cortex. *Curr Opin Neurobiol* 2008; **18**:28-35.

72 Sugitani Y, Nakai S, Minowa O, *et al*. Brn-1 and Brn-2 share crucial roles in the production and positioning of mouse neocortical neurons. *Genes Dev* 2002; **16**:1760-1765.

73 Pang ZP, Yang N, Vierbuchen T, *et al*. Induction of human neuronal cells by defined transcription factors. *Nature* 2011; **476**:220-223.

74 Yoo AS, Sun AX, Li L, *et al*. MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature* 2011; **476**:228-231.

75 Pfisterer U, Kirkeby A, Torper O, *et al*. Direct conversion of human fibroblasts to dopaminergic neurons. *Proc Natl Aca Sci USA* 2011; **108**:10343-10348.

76 Caiazzo M, Dell'Anno MT, Dvoretskova E, *et al*. Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* 2011; **476**:224-227.

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

78 Qiang L, Fujita R, Yamashita T, *et al*. Directed conversion of Alzheimer's disease patient skin fibroblasts into functional neurons. *Cell* 2011; **146**:359-371.

79 Huang P, He Z, Ji S, *et al*. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* 2011; **475**:386-389.

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

81 Zaret KS, Grompe M. Generation and regeneration of cells of the liver and pancreas. *Science* 2008; **322**:1490-1494.

82 Son EY, Ichida JK, Wainger BJ, *et al*. Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell* 2011; **9**:205-218.

83 Ambasudhan R, Talantova M, Coleman R, *et al*. Direct reprogramming of adult human fibroblasts to functional neurons under defined conditions. *Cell Stem Cell* 2011; **9**:113-118.

84 Szabo E, Rampalli S, Risueño RM, *et al*. Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* 2010; **468**:521-526.

85 Cobaleda C, Jochum W, Busslinger M. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature* 2007; **449**:473-477.

86 Laiosa CV, Stadtfeld M, Xie H, de Andres-Aguayo L, Graf T. Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBP alpha and PU.1 transcription factors. *Immunity* 2006; **25**:731-744.

87 Heinrich C, Blum R, Gascón S, *et al*. Directing astroglia from the cerebral cortex into subtype specific functional neurons. *PLoS Biol* 2010; **8**:e1000373.

88 Berninger B, Costa MR, Koch U, *et al*. Functional properties of neurons derived from *in vitro* reprogrammed postnatal astroglia. *J Neurosci* 2007; **27**:8654-8664.

89 Marro S, Pang ZP, Yang N, *et al*. Direct lineage conversion of terminally differentiated hepatocytes to functional neurons. *Cell Stem Cell* 2011; **9**:374-382.

90 Ferber S, Halkin A, Cohen H, *et al*. Pancreatic and duodenal homeobox gene 1 induces expression of insulin genes in liver and ameliorates streptozotocin-induced hyperglycemia. *Nat Med* 2000; **6**:568-572.

91 Kaneto H, Nakatani Y, Miyatsuka T, *et al*. PDX-1/VP16 fusion protein, together with NeuroD or Ngn3, markedly induces insulin gene transcription and ameliorates glucose tolerance. *Diabetes* 2005; **54**:1009-1022.

92 Miyatsuka T, Kaneto H, Kajimoto Y, *et al*. Ectopically expressed PDX-1 in liver initiates endocrine and exocrine pancreas differentiation but causes dysmorphogenesis. *Biochem Biophys Res Commun* 2003; **310**:1017-1025.

93 Sapir T, Shternhall K, Meivar-Levy I, *et al*. *Cell*-replacement therapy for diabetes: Generating functional insulin-producing tissue from adult human liver cells. *Proc Natl Aca Sci USA* 2005; **102**:7964-7969.

94 Kojima H, Fujimiya M, Matsumura K, *et al*. NeuroD-betacellulin gene therapy induces islet neogenesis in the liver and reverses diabetes in mice. *Nat Med* 2003; **9**:596-603.

95 Hanna J, Saha K, Pando B, *et al*. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* 2009; **462**:595-601.

96 Egli D, Birkhoff G, Eggan K. Mediators of reprogramming: transcription factors and transitions through mitosis. *Nat Rev Mol Cell Biol* 2008; **9**:505-516.

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

98 Cirillo LA, Zaret KS. An early developmental transcription factor complex that is more stable on nucleosome core particles than on free DNA. *Mol Cell* 1999; **4**:961-969.

99 Davidson EH, Levine MS. Properties of developmental gene regulatory networks. *Proc Natl Aca Sci USA* 2008; **105**:20063-20066.

100 Taberlay PC, Kelly TK, Liu CC, *et al*. Polycomb-repressed genes have permissive enhancers that initiate reprogramming. *Cell* 2011; **147**:1283-1294.

101 Gal-Yam EN, Egger G, Iniguez L, *et al*. Frequent switching of Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. *Proc Natl Aca Sci USA* 2008; **105**:12979-12984.

102 Azuara V, Perry P, Sauer S, *et al*. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 2006; **8**:532-538.

103 Bernstein BE, Mikkelsen TS, Xie X, *et al*. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 2006; **125**:315-326.

104 Kim JB, Zaehres H, Wu G, *et al*. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* 2008; **454**:646-650.

105 Maherali N, Ahfeldt T, Rigamonti A, Utikal J, Cowan C,

Hochedlinger K. A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* 2008; **3**:340-345.

106 Aasen T, Raya A, Barrero MJ, *et al*. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol* 2008; **26**:1276-1284.

107 Ng RK, Gurdon JB. Epigenetic memory of active gene transcription is inherited through somatic cell nuclear transfer. *Proc Natl Aca Sci USA* 2005; **102**:1957-1962.

108 Ng RK, Gurdon JB. Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. *Nat Cell Biol* 2008; **10**:102-109.

109 Kim K, Doi A, Wen B, *et al*. Epigenetic memory in induced pluripotent stem cells. *Nature* 2010; **467**:285-290.

110 Kim K, Zhao R, Doi A, *et al*. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat Biotechnol* 2011; **29**:1117-1119.

111 Polo JM, Liu S, Figueroa ME, *et al*. *Cell* type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat Biotechnol* 2010; **28**:848-855.

112 Davidson EH, Erwin DH. Gene regulatory networks and the evolution of animal body plans. *Science* 2006; **311**:796-800.

113 Gaetz J, Clift KL, Fernandes CJ, *et al*. Evidence for a critical role of gene occlusion in cell fate restriction. *Cell Res* 2012; **22**:848-858.

114 Miller RA, Ruddle FH. Pluripotent teratocarcinoma-thymus somatic cell hybrids. *Cell* 1976; **9**:45-55.

115 Mak W, Nesterova TB, de Napoles M, *et al*. Reactivation of the paternal X chromosome in early mouse embryos. *Science* 2004; **303**:666-669.

116 Efe JA, Hilcove S, Kim J, *et al*. Conversion of mouse fibroblasts into cardiomyocytes using a direct reprogramming strategy. *Nat Cell Biol* 2011; **13**:215-222.

117 Meissner A, Wernig M, Jaenisch R. Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nat Biotechnol* 2007; **25**:1177-1181.

118 Nakagawa M, Koyanagi M, Tanabe K, *et al*. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* 2008; **26**:101-106.

119 Wernig M, Meissner A, Cassady JP, Jaenisch R. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell* Stem *Cell* 2008; **2**:10-12.

120 Christen B, Robles V, Raya M, Paramonov I, Izpisúa Belmonte JC. Regeneration and reprogramming compared. *BMC Biol* 2010; **8**:5.

121 Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 2000; **24**:372-376.

122 Zeineddine D, Papadimou E, Chebli K, *et al*. Oct-3/4 dose dependently regulates specification of embryonic stem cells toward a cardiac lineage and early heart development. *Dev Cell* 2006; **11**:535-546.

123 Boyer LA, Lee TI, Cole MF, *et al*. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 2005; **122**:947-956.

124 Kopp JL, Ormsbee BD, Desler M, Rizzino A. Small increas-

es in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cell*s 2008; **26**:903-911.

125 Zhao S, Nichols J, Smith AG, Li M. SoxB transcription factors specify neuroectodermal lineage choice in ES cells. *Mol Cell Neurosci* 2004; **27**:332-342.

126 Kim J, Efe JA, Zhu S, *et al*. Direct reprogramming of mouse fibroblasts to neural progenitors. *Proc Natl Aca Sci USA* 2011; **108**:7838-7843.

127 Thier M, Wörsdörfer P, Lakes YB, *et al*. Direct conversion of fibroblasts into stably expandable neural stem cells. *Cell Stem Cell* 2012; **10**:473-479.

128 Hiramatsu K, Sasagawa S, Outani H, Nakagawa K, Yoshikawa H, Tsumaki N. Generation of hyaline cartilaginous tissue from mouse adult dermal fibroblast culture by defined factors. *J Clin Invest* 2011; **121**:640-657.

129 Emslie D, D'Costa K, Hasbold J, *et al*. Oct2 enhances antibody-secreting cell differentiation through regulation of IL-5 receptor alpha chain expression on activated B cells. *J Exp Med* 2008; **205**:409-421.

130 Lengner CJ, Camargo FD, Hochedlinger K, *et al*. Oct4 expression is not required for mouse somatic stem cell self-renewal. *Cell Stem Cell* 2007; **1**:403-415.

131 Hochedlinger K, Yamada Y, Beard C, Jaenisch R. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* 2005; **121**:465-477.

132 Samavarchi-Tehrani P, Golipour A, David L, *et al*. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 2010; **7**:64-77.

133 Mikkelsen TS, Hanna J, Zhang X, *et al*. Dissecting direct reprogramming through integrative genomic analysis. *Nature* 2008; **454**:49-55.

134 Boiani M, Eckardt S, Schöler HR, McLaughlin KJ. Oct4 distribution and level in mouse clones: consequences for pluripotency. *Genes Dev* 2002; **16**:1209-1219.

135 Buganim Y, Faddah DA, Cheng AW, *et al*. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 2012; **150**:1209-1222.

136 Zappone MV, Galli R, Catena R, *et al*. Sox2 regulatory sequences direct expression of a (beta)-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells. *Development* 2000; **127**:2367-2382.

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

138 Polo JM, Anderssen E, Walsh RM, *et al*. A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* 2012. *In press*

139 Kurian L, Sancho-Martinez I, Nivet E, *et al*. Conversion of human fibroblasts to angioblast-like progenitor cells. *Nat Meth* 2012 Dec 2. doi:10.1038/nmeth.2255