Ann. N.Y. Acad. Sci. ISSN 0077-8923

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Issue: *Hematopoietic Stem Cells VIII*

Reprogramming cell fates: insights from combinatorial approaches

Carlos-Filipe Pereira, Ihor R. Lemischka, and Kateri Moore

Department of Developmental and Regenerative Biology, Black Family Stem Cell Institute, Mount Sinai School of Medicine, New York, New York

Address for correspondence: Carlos-Filipe Pereira and Kateri Moore, Department of Developmental and Regenerative Biology, Mount Sinai School of Medicine, 1 Gustave L. Levy Place, Box 1496, New York, NY 10029. filipe.pereira@mssm.edu and Kateri.Moore@mssm.edu

Epigenetic reprogramming can be achieved in different ways, including nuclear transfer, cell fusion, or the expression of transcription factors (TFs). Combinatorial overexpression provides an opportunity to define the minimal core network of TFs that instructs specific cell fates. This approach has been employed to induce mouse and human pluripotency and differentiated cell types from cells that can be also as distant as cells from different germ layers. This suggests the possibility that any specific cell type may be directly converted into another if the appropriate reprogramming TF core is determined. Herein, we review the factors used for reprogramming multiple cell identities and raise the question of whether there is a common underlying blueprint for reprogramming factors. In addition to the generation of human cell types of interest for cell-replacement therapies, we propose that the TF-mediated conversion of differentiated cell types, especially somatic stem cells, will have an impact on our understanding of their biological development.

Keywords: reprogramming; transcription factor; induced pluripotency; transdetermination; iPS; direct conversion

Introduction

Once committed, the differentiated state of a cell is normally stable and can be inherited through cell division. Under certain conditions, cell fate can, however, be modified or reversed.[1–3] Epigenetic reprogramming can be achieved experimentally in different ways, including through nuclear transfer, cell fusion, or the forced expression of transcription factors (TFs).[4]

The first evidence of successful nuclear reprogramming of somatic cells was provided by the live birth of animals cloned by injection of differentiated somatic nuclei into eggs.[5] This demonstrates that terminally differentiated cells retain cell plasticity and can be reprogrammed to produce an adult cloned animal.[5,6] It remained a formal possibility, however, that the donor nuclei that gave rise to the rarely observed clones were actually derived from tissue stem cells, which represent a small proportion of adult tissues. Subsequently, the successful generation of cloned mice from lymphocytes[7,8] unambiguously demonstrated that terminal differentiation does not restrict the potential of the nucleus to support development. Importantly, this established the principle that mechanisms underlying lineage restriction and cell identity are ultimately reversible.

The combination of two cell types through cell fusion to form somatic cell hybrids and heterokaryons yielded evidence that mammalian gene expression can be altered by diffusible *trans*-acting factors.[9] Cells derived from mesoderm, endoderm, and ectoderm could be dominantly reprogrammed into mouse myotubes that express muscle specific genes.[10] Cells that were more closely related to muscle—mesodermal derivatives—consistently expressed muscle genes sooner and to a greater extent than the cells of ectodermal or endodermal origin.[10] These studies established that differentiated cells do retain flexible lineage potential and that lineage conversion and gene activation can occur in absence of DNA replication.[11] The dominance of pluripotent cells over differentiated cells has also been shown in experimental hybrids and heterokaryons

doi: 10.1111/j.1749-6632.2012.06508.x

made between somatic cells and embryonic stem cells (ESCs).[12–14] The *trans*-acting factors mediating dominant reprogramming may reside in the nucleus or in the cytoplasm. This question was initially addressed by separating the nuclear compartment (karyoplast) from the cytoplasmic compartment (cytoplast) of an ESC; these elements were then individually fused with neuronal cells.[15] Only karyoplasts were able to reprogram upon cell fusion, suggesting that nuclear factors are essential for reprogramming. This conclusion is consistent with cloning experiments in amphibians[16] and mice,[17] which indicate that successful reprogramming depends on direct injection of nuclei into the germinal vesicle or into a metaphase oocyte, where nuclear factors are available in the cytoplasm. Interestingly, the overexpression of a single TF in somatic cells was unexpectedly found to modify or override lineage outcome, first in *Drosophila melanogaster*[18] and subsequently in mammals,[19] leading to remarkable changes in cell fate.

In 2006, Takahashi and Yamanaka reported the breakthrough discovery that a combination of defined TFs was sufficient to reprogram several somatic cell types to pluripotency (termed "iPS" for induced pluripotent stem).[20–22] In the initial study,[20] the authors sought to identify genes expressed in ESCs that would be sufficient to induce pluripotency. With a pioneer approach, 24 candidate genes were simultaneously expressed in fibroblasts using retroviral transduction. This resulted in the reprogramming of a small percentage of fibroblasts toward iPS cells. Genes were then systematically removed from the cocktail to define the minimal and optimal combinations of factors for inducing pluripotency. The same rationale has been more recently applied to induce differentiated cell types from cells that can be as distant as cells from different germ layers. Reprogrammed cells, by direct lineage conversion, include macrophages, brown fat cells, cardiomyocytes, neurons, hepatocytes, and beta cells.[23–28] Cell fate conversion to each of these cell types was accomplished by expression of a signature group of two to five TFs in unrelated cell types. Recently, the expression of miRNAs was also shown to either reprogram cells to pluripotency[29] or to assist in direct conversion toward the neural lineage in combination with TFs.[30] Because miRNAs generally target hundreds of mRNAs that coordinate expression of many different proteins, the underlying mechanism may involve the indirect activation of TFs. Collectively, these findings opened an avenue for combinatorial overexpression screening to define the minimal TF core required to induce any cell type of interest. Can this approach be applied to induce somatic stem cells? Is the induction of multipotency fundamentally different from direct lineage conversions? Do TFs share similar biochemical properties that confer the ability to trigger cellular reprogramming? Here, we review these questions with a focus on the *reprogramming factors*, TFs that are sufficient to instruct cell type conversion when overexpressed in unrelated somatic cells. We compiled data from TF overexpression studies that lead to induction of cell types and analyzed the TFs according to several features that may facilitate the selection of factors to induce additional cell fates. The establishment of cell type–specific TF minimal networks will make possible the direct generation of patient-specific cell types of interest for cell replacement or disease modeling. In addition, it also provides a novel approach to interrogate aspects of cell fate, stem cell function, and cellular differentiation.

Reprogramming cell fates by combinatorial expression of transcription factors

The induction of pluripotent stem cells was successfully achieved from mouse embryonic and adult fibroblasts after viral-mediated transduction of *Oct4*, *Sox2*, *c-Myc*, and *Klf4*. Importantly, iPS cells generated postnatal chimeras, contributed to the germ line,[31–33] and generated late-gestation embryos through tetraploid complementation[31] (the most stringent test for developmental potency). In addition, iPS cells were shown to be transcriptionally and epigenetically similar to ES cells. iPS cells showed reactivation of the somatically silenced X-chromosome, DNA demethylation of pluripotency-associated genes, and acquisition of chromatin structure that closely resemble ES cells.[31–33] Expression of the reprogramming factors in fibroblasts appears to initiate a cascade of events that leads to the conversion of a small percentage of cells. Similar to what is observed in nuclear transfer,[34] iPS reprogramming appears to be influenced by donor cell type. The same four TFs have been shown to drive mouse hepatocytes, stomach cells,[21] pancreatic beta cells,[35] and immature B lymphocytes to iPS, whereas adult mouse B lymphocytes[36]

required silencing of the B cell–specific TF *Pax5* or the additional expression of *CEBPα*. Conversely, just two factors (*Oct4* and *Klf4*) or one factor (*Oct4*) were sufficient to reprogram adult neural stem cells and dermal papilla cells, which endogenously express *Sox2* and *c-Myc*.[37–41] In 2007, reprogramming of human fibroblasts was achieved in parallel approaches using *Oct4*, *Sox2,* and either *Nanog* plus *Lin28*[42] or *Klf4* plus *c-Myc*.[22,43,44] Lin28 is the only protein of the combination that is not a TF. Human cell types converted to iPS include fetal fibroblasts, adult dermal fibroblasts, bone marrow–derived mesenchymal cells, and keratinocytes. In order to reprogram keratinocytes a combination of five factors (*Oct4*, *Sox2, Klf4, c-Myc,* and *Nanog*) was used.[45]

Several combinations of factors have been used for iPS induction in mice and humans (Table 1, upper left panel), affecting mainly the efficiency and rapidity of reprogramming (reviewed in Ref. 46). iPS colonies were derived from mouse fibroblasts using the combination of *Oct4*, *Sox2*, and *Klf4*.[47] Similarly, human iPS can be generated with either *Oct4*, *Sox2*, and *Klf4* (no *c-Myc*) or *Oct4*, *Sox2*, and *Nanog* (no *Lin28)*.[42,47] TFs that belong to the same family and are closely related can replace the function of some of the reprogramming factors; *Sox2* can be replaced by *Sox1* and *Sox3*, *Klf4* by *Klf2* and *Klf5*, and *c-Myc* by *N-Myc* and *L-Myc*.[42,47,48] TFs that belong to other families can also replace reprogramming factors, for example, the orphan nuclear receptor *Esrrb* can replace the function of both *Klf4* and *c-Myc*.[49] Interestingly, *Oct4* cannot be replaced by the closest homologues *Oct1* or *Oct6*.[47] Indeed, the only factor that has been reported to replace *Oct4* function in reprogramming is the nuclear receptor *Nr5a2*, previously shown to regulate *Oct4* levels.[50,51] These observations suggest that *Oct4* occupies an irreplaceable, yet mechanistically unidentified, role in induced pluripotency. This is supported by the observation that ESCs conditionally depleted for *Oct4* (before differentiation or loss of other reprogramming factors) lose dominant reprogramming capacity in heterokaryons, in contrast to *Sox2*.[13] Chromatin remodeling factors (*Brg1*, *Baf155*),[52] histone demethylation (*Jhdm1a/1b*),[53] nuclear receptors (*RAR-gamma*, *Nr5a2*),[54] additional TFs (*Tbx3, Glis1, Prdm14*),[55–57] and co-activators (*Utf1*),[58] have been shown to increase in the efficiency and rapidity of reprogramming when added in combination with *Oct4*, *Sox2*, *Klf4* (± *c-Myc*). Similar outcomes were achieved by ablation of factors that impede reprogramming: silencing of p53 and other regulators of senescence (acting as a major barrier to reprogramming by limiting cell cycle and inducing apoptosis[59]), repression of lineage-specifying TFs and inhibiting DNA methylation activity.[60,61] Together these reprogramming experiments suggest that iPS-inducing factors can be separated into two groups: core reprogramming factors that play a central role for induced pluripotency, including irreplaceable TFs (*Oct4*) and necessary TFs but interchangeable for closely related or functionally equivalent factors (*Sox2*, *Klf4, Nanog,* and *Esrrb*); and pro-programming factors,[57] which facilitate the reprogramming process resulting in increased efficiency and speed. The latter set can act on multiple cellular pathways such as cell cycle, chromatin remodeling, DNA demethylation, and mesenchymal-to-epithelial transition.[62–64]

It is also possible to induce cells directly into other cell fates using the appropriate factors (Table 1), a process that has been named transdifferentiation or direct conversion.[65] Conversion of fibroblasts into myoblasts by expressing *MyoD* was first demonstrated in mammals in 1987 by Davis and Weintraub.[19] However, ectopic expression of *MyoD* in ectoderm lineages *in vivo*[66] or in hepatocytes *in vitro*[67] does not initiate the complete conversion of these lineages into differentiated muscle. This suggests that *MyoD*, although a central player in the skeletal muscle program, needs cooperation from additional, yet unidentified factors to convert more distant cell types. Similarly, the TFs *CEBPα* and *CEBPβ* can induce committed mature B lymphocytes to become macrophages,[68] but an additional factor (*PU.1*) is required to convert fibroblasts.[23] Forced coexpression of *Prdm16* and *CEBPβ* was sufficient to convert mouse and human dermal fibroblasts into functional brown fat-like cells.[24]

More recently, the combinatorial overexpression approach was used to screen for induction of cell types emanating from different germ layers. After testing an initial pool of 19 neuronal lineage-specific TFs, a combination of three factors (*Ascl1*, *Brn2,* and *Mytl1*) was sufficient to induce neurons from fibroblasts, with almost 20% efficiency in less than one week.[26] The conversion of human fibroblasts to fully functional neurons appears to be greatly aided by the helix-loop-helix TFs *NeuroD1* and

Table 1. Reprogramming factors and their features

Factors to induce:	Mouse	Human	SS-TF	TF family/domains	Loss of function	Cell proliferation	Pioneer activity
Pluripotency	20 20 45 49 52 57 53 51 54 55	22 22 42 42 57 58 56					
Oct-4			Yes	Pou/Homeodomain	+	-	ND
Sox2			Yes	HMG-box/SOXp	+	-	ND
Klf4			Yes	Zn-finger	-*	+	ND
c-Myc			Yes	Helix-loop-helix/ Myc-N/Myc-LZ	+	+	ND
Esrrb			Yes	NR Zn-finger	+	-	ND
Nanog			Yes	Homeodomain	+	-	ND
Lin28			No	RNA-binding	-	-	ND
Tbx3			Yes	T-box	+	+	ND
Nr5a2			Yes	NR Zn-finger	+	+	ND
RAR-gamma			Yes	NR Zn-finger	-	-	ND
Utf1			No	Co-activator	-	-	ND
Brg1			No	Chromatin	+	-	ND
Baf155			No	remodelling		-	ND
Glis1			Yes	Zn-finger	ND	-	ND
Jhdm1a/1b			No	Histone demethylase	ND	-	ND
Prdm14			Yes	Zn-finger, SET	+	-	ND
Myocytes	19	97					
MyoD			Yes	Helix-loop-helix	-*	-	ND
Cardiomyocytes	25						
Tbx5		ND	Yes	T-box	+	-	ND
Mef2c			Yes	MADS	+	-	ND
Gata4			Yes	Zn-finger, Gata-type	+	-	ND
Neurons	26 70	69 71 70					
Ascl1	G D	G D D	Yes	Helix-loop-helix	+	-	ND
Brn2			Yes	Pou/Homeodomain	+	-	ND
Myt1l			Yes	Zn-finger	ND	-	ND
NeuroD1/2			Yes	Helix-loop-helix	+	-	ND
Lmx1a			Yes	Homeodomain	+	-	ND
Nurr1			Yes	NR Zn-finger	+	-	ND
Foxa2			Yes	Forkhead	+	-	+
Hepatocytes	27 72						
Gata4			Yes	Zn-finger, Gata-type	+	-	+
Hnf1α			Yes	Homeobox	+	-	ND
Hnf4α		ND	Yes	NR Zn-finger	+	-	ND
Foxa3			Yes	Forkhead	-*	-	+
Foxa2			Yes	Forkhead	+*	-	+
Foxa1			Yes	Forkhead		-	+
Macrophages	23, 68						
PU.1		ND	Yes	Ets	+	-	+
CEBPα/β			Yes	bZIP	+	-	ND
Brown fat cells	24	24					
Prdm16			Yes	Zn-finger, SET	+	-	ND
CEBPβ			Yes	bZIP	+	-	ND
Beta cells	28 28						
Ngn3			Yes	Helix-loop-helix	+	-	ND
NeuroD1		ND	Yes	Helix-loop-helix	+	-	ND
Pdx1			Yes	Homeobox	+	-	ND
Mafa			Yes	bZIP	+	-	ND

TF – Transcription factor. SS-TF – Sequence-specific transcription factor. NR – nuclear receptor. ND – not determined. G – Glutamatergic neurons. D – Dopaminergic neurons. *Reported compensatory effects by homologous TFs.

NeuroD2.[69] Two recent studies have reported that the addition of TFs that are involved in dopaminergic neuronal development can generate dopaminergic neurons from human and mouse fibroblasts that express tyrosine hydroxylase and release dopamine. Each study used a distinct group of TFs for conversion (*Ascl1, Brn2*, *Myt1l* + *Lmx1a,* and *Foxa2* versus *Ascl1*, *Nurr1,* and *Lmx1a*).[70,71] Interestingly, transduction of *Ascl1* alone was sufficient to induce some neural traits in fibroblasts, such as the expression of pan-neuronal proteins. This gene was present in all neural reprogramming cocktails and may be analogous to *Oct4* for the induction of pluripotency and *MyoD* for the induction of muscle. Leda *et al.* used an iterative process of elimination to define a minimal pool of three cardiac-specific TFs (*Gata4*, *Mef2c*, *Tbx5*) that directly induced mouse fibroblast transdifferentiation into cardiomyocytes.[25] Similar to the induction of neurons, the conversion of cardiomyocytes was rapid (one week) and efficient (>6% αMHC + cTnT + cells). Recently, two groups have demonstrated the conversion of mouse fibroblasts to

hepatocyte-like cells by forced expression of *Gata4*, *Hnf1a,* and *Foxa3* and inactivation of $P19^{Arf}$ or by expression of *Hnf4*α and *Foxa1*, *Foxa2,* or *Foxa3*.[27,72] Finally using a strategy of re-expressing key developmental regulators *in vivo*, Melton and colleagues identified a specific combination of three TFs (*Ngn3*, *Pdx1,* and *Mafa*) that converts differentiated pancreatic exocrine cells in adult mice into cells that closely resemble beta cells.[28] Interestingly, attempts to reprogram skeletal muscle and embryonic fibroblasts (mesodermal cell types) using this combination were unsuccessful.

In addition to these direct reprogramming studies, some reports suggest that overexpression of pluripotency-reprogramming factors in permissive culture conditions can generate cell types other than iPS. For example, the induction of hematopoietic progenitors was achieved by overexpression of *Oct4* in the presence of hematopoietic cytokines[73] and the induction of cardiomyocytes and neural progenitors by the induction of the *Oct4*, *Sox2*, *Klf4,* and *c-Myc*.[74,75] These results are unexpected because the generated cell types do not express the TFs used for induction. Future experiments may address the mechanisms underlying these reprogramming events that may involve transit through a partially reprogrammed state further instructed with culture conditions to differentiate into specific cell types.

In summary, seven cell types representative of the three germ layers were induced by overexpression of TF cocktails. With this approach, master regulators of cell fate induction have been allocated to these cell types. Reassuringly, the majority of these genes are essential for the generation of the target cell types during development. It remains an open question whether there are common features between reprogramming-competent TFs. Therefore, it is important to pay careful attention to those proteins and their molecular function.

Transcription factors with reprogramming abilities—are there common features?

Diverse arrays of proteins are crucial for successful transcription by RNA polymerase in eukaryotic cells. These proteins include general TFs, cofactors (coactivators and corepressors) and chromatin modifying proteins. In addition, sequence-specific DNA-binding TFs direct transcription initiation to specific promoters. The vast majority (>80%) of the reported reprogramming factors (Table 1) are annotated as sequence-specific TFs.[76] The exceptions include the RNA-binding protein Lin28, the chromatin remodelers Brg1 and Baf155, the histone demethylases Jhmd1a/1b and the coactivator Utf1. Interestingly, the exclusion of those factors from the reprogramming combinations only reduces the efficiency but does not abolish conversion, placing those genes in the pro-programming group and not in the minimal core of TFs to instruct cellular identity. DNA methyltransferases, chromatin remodelers, and chromatin modifiers were included in the initial pool of several combinatorial screens,[20,25,27] but none of those molecules were incorporated in the minimal combination to induce cell fates. Chromatin remodeling and DNA demethylation was shown to be required for converting the epigenome of one cell type to another.[60,61,77] However, with combinatorial overexpression approaches these molecules would lack the capacity to instruct. We suggest that factors with reprogramming abilities have proven or predicted sequence-specific DNA-binding TF activity.

There are currently 1,866 human proteins predicted to have sequence-specific TF activity in the gene ontology database.[76] A bioinformatic analysis of the TFs of the human genome[78] showed that in each tissue 150–300 TFs are expressed, and only a proportion (<35%) show a high degree of tissue-specificity (high expression restricted to up to three tissues). We used the publicly available BIOGPS database (www.biogps.org; GeneAtlas MOE430) to assess the expression profile of the known reprogramming factors across many tissues and cell lines (Fig. 1). The available gene expression data for the reprogramming factors to induce pluripotency, myoblasts, cardiomyocytes, neurons, hepatocytes, and macrophages were extracted from the database and hierarchical clustering was performed. The pattern of gene expression of these genes was sufficient to clearly cluster ESCs, muscle, heart, brain, liver, and macrophages. Interestingly, reprogramming factors show very restricted expression patterns in the cell type/tissue that they induce, especially *Oct4*, *MyoD*, *Tbx5*, and *Ascl1*. This implies that in addition to the comparison of the TF-transcriptome in the initial and target cell types, including expression in other cell types from other tissues may greatly increase the specificity of the analysis. Online databases such as BIOGPS or Unigene Differential Digital Display (www.ncbi.nlm.nih.gov/UniGene/ddd.cgi)

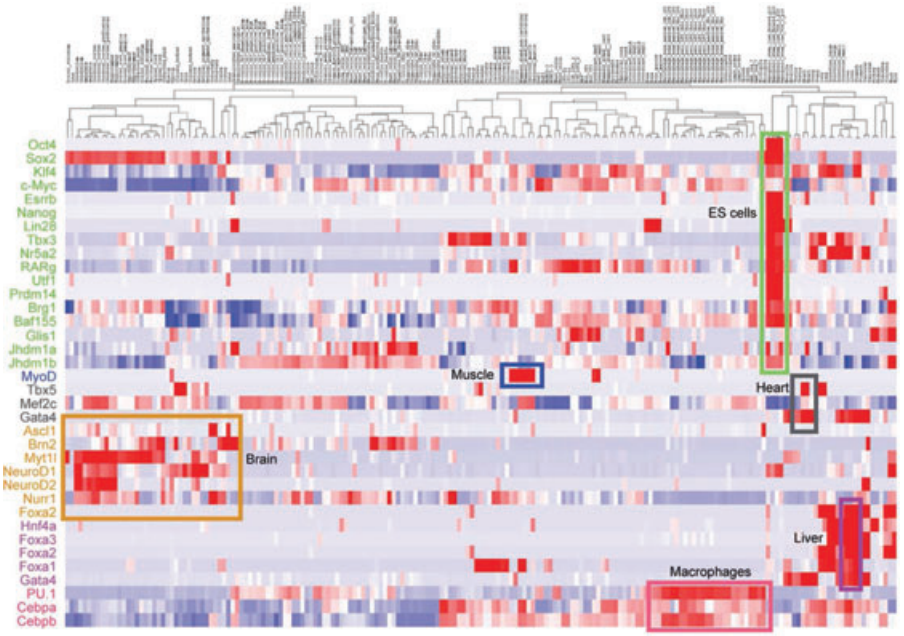

Figure 1. **Heatmap representation of reprogramming factor expression across multiple tissues. Heatmap showing the gene expression of transcription factors (rows) in multiple tissues and cell lines (columns). Intersecting cells are shaded according to expression level (dark red for high expression and blue for low expression relative to the mean). Tissues are grouped according to hierarchical clustering based on the gene expression of 35 transcription factors. The boxes highlight high expression of reprogramming factors in induced tissue/cell types and the lack of broad expression on other tissues. Green, induced pluripotency; blue, induced myoblasts; gray, induced cardiomyocytes; brown, induced neurons; purple, induced hepatocytes; pink, induced macrophages. These data were extracted from the public online database BIOGPS (www.biogps.org; GeneAtlas MOE430).**

provide good tools for broad gene expression comparisons.[20] This kind of analysis may aid the identification of lineage-restricted reprogramming factors.

The implication that a small set of lineage-restricted sequence-specific TFs will instruct cell fates predicts that: loss of function of those genes would affect the generation of the cell type during development, and gene expression may be tightly controlled in other cell types. We used the MGI database (www.informatics.jax.org) and PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) to interrogate the loss of function phenotype (knockout or knockdown approaches) of the reprogramming factor deletion during tissue development/target cell type (Table 1). Interestingly the majority of the factors used for reprogramming have an impact either on the generation or maintenance of the target cell type (Table 1). The exceptions include factors whose function is compensated by homologous TFs on the *Klf* , *MyoD,* and the *Foxa* gene-families.[79] For example, *Klf4* depletion does not affect ESC self-renewal; however, simultaneous depletion of *Klf2*, *Klf4,* and *Klf5* leads to ESC differentiation.[80] Similarly *MyoD*-null mice are viable,[81] but when *Myf5*, another helix-loop-helix TF is deleted in combination with *MyoD*, there is a total absence of muscle fibers and precursor myoblasts.[82] In summary, reprogramming factors may lie at the intersection of sequence-specific TFs, factors that influence the generation of a target cell type during development, and tissue/cell type-restricted gene expression. These criteria can be used to rank candidates for combinatorial overexpression screens to determine TF core networks to induce cell fates. Further experimentation will be required to address the mechanisms of silencing of those genes in unrelated cell types and how they may differ from other TFs. In this regard, *Ascl1* is controlled not only by repressive histone modifications in ESCs but also replicates late during S-phase and associates with the nuclear periphery (temporal and spatial constraints associated with gene repression). Upon neural differentiation the *Ascl1* gene shifts to replicate early in S-phase

and translocates to the nuclear interior where it is expressed.[83]

It is intriguing that some TFs are included in several different lineage-specific cocktails (Table 1). These factors include *NeuroD1* (neurons and beta-cells), *CEBPβ* (Brown fat cells and macrophages), *Gata4* (cardiomyocytes and hepatocytes), and *Foxa2* (hepatocytes and neurons). *Glis1* also enhanced the expression of *Foxa2*, indicating that this TF may also have a role in the promotion of iPS cell generation possibly by antagonizing the epithelial to mesenchymal transition.[57] This suggests that a small number of factors among the horde of sequence-specific TFs may have unique reprogramming features. For example, reprogramming factors may act as "pioneer factors," TFs found to be able to access their DNA target sites in silent chromatin when other factors cannot and continue to access the DNA prior to the time of transcriptional activation (reviewed by Refs. 84 and 85). Among the list of reprogramming factors (Table 1), five factors (*Gata4*, *Foxa1*, *Foxa2*, *Foxa3*, *PU.1*) have been implicated as having pioneer activity. In undifferentiated endoderm cells, the Foxa and Gata TFs are among the first to engage silent genes, such as the *Alb1* gene, helping to endow competence for cell type specification. Foxa proteins can bind their target sites in highly compacted chromatin and open up the local region for other factors to bind. In addition, Foxa1 protein remains bound to chromatin during mitosis.[86] PU.1 itself can expand the linker region between nucleosomes and promote local histone modifications, likely contributing to its ability to enhance binding of other factors.[87,88] It would be interesting to test the pioneer TF model and attempt to reprogram fibroblasts to beta cells by incorporating *Gata4*/*Foxa2* in the reported TF-core of *Pdx1*, *Mafa,* and *Ngn3*/*NeuroD1*.

Future experiments will be required to address whether there is a common underlying blueprint for reprogramming factors. Genome-wide location studies, *in vivo* footprinting, and determining their interaction partners may begin to shed light on the molecular function of these genes throughout various stages of development. This will also allow us to start addressing the mechanistic logic of TF-induced reprogramming by building the network of regulatory interactions. An additional important question is whether inducing pluripotent stem cells is fundamentally a different process from direct lineage conversion. Expanding combinatorial approaches to reprogram other cell identities, including somatic stem cells, will be crucial to define the nature of reprogramming.

Features of reprogramming toward pluripotency and somatic cells—perspectives to induce somatic stem cells

A comparison of the iPS cell reprogramming and direct lineage conversions reveal several distinct

Table 2. Reprogramming to pluripotency, somatic cells, and potentially somatic stem cells

	iPS cell reprogramming	Direct conversion of somatic cells	Somatic stem cell reprogramming
Target cell–type features			
Pluripotency	+	–	–
Multipotency	–	–	+
Self-renewal	+	–	+
Reprogramming features			
Reprogramming factors that regulate cell cycle	+	–	?
Reprogramming efficiency	Low	High	?
Reprogramming dynamics	Slow	Fast	?
Potential tumor risk	Low	High	?
Target cell generation	Difficult[a]	Easy	?
Cell scaling	Feasible	Limited	?
Gene correction	Feasible	Limited	?

[a]Protocols are lacking to efficiently generate some cell types, for example, hematopoietic stem cells.

features, although both involve TFs to trigger reprogramming (Table 2). Somatic lineage conversions are very rapid, with the first lineage reporters of the target cell type expressed just days after gene induction.[25,26] Inducing iPS cells typically takes 10–20 days and occurs with much lower efficiency.[20,22] Despite the early initiation of conversion, the activation of markers of mature cells, and the acquisition of functional properties appears to be delayed and to continue for several weeks; a process analogous to differentiation. Nonetheless, lineage conversion seems to take place without the generation of a progenitor cell type.[25,26,89] TF-mediated lineage reprogramming seems to be a direct phenotypic induction and not dedifferentiation followed by differentiation to an alternative fate.

Some direct cell fate conversion examples do not require cell division.[26,90] Conversely, cell proliferation increases the efficiency of generation of iPS,[91] and the reprogramming process is significantly delayed and less efficient in the absence of the *c-Myc* oncogene. We interrogated whether those differences are translated in the properties of TFs used for reprogramming, using gene ontology[76] to address their molecular function (Table 1). Interestingly, the factors that have been implicated in the control of the cell cycle or inducing proliferation are restricted to the induction of pluripotency (Tables 1 and 2). This can be either a feature of induced pluripotency or a broader determinant of self-renewal. The latter may have implications on future attempts to induce somatic stem cells, which also self-renew (Table 2). If this prediction is true, we would expect that to induce somatic stem cells, such as hematopoietic stem cells (HSCs), additional factors to induce self-renewal would be required in addition to the lineage-restricted reprogramming factors. On the other hand, triggering senescence/cell-cycle arrest has been shown to be a major barrier for iPS-reprogramming and also for the direct conversion to hepatocytes.[27,59]

One of the most pressing medical objectives is to obtain patient-specific cell types for cell replacement therapy. Although iPS cells have the clear advantage of unlimited growth and scalability, the potential risk for tumor formation is high. Directly converted somatic cells would be presumably less tumorigenic, but obtaining large numbers of cells and gene correction strategies will be challenging (Table 2). In addition, the differentiation of human ESCs and iPS cells is highly variable, cell-line dependent, and generates immature cells that differ from those found in mature organs *in vivo*.[92] Using the current protocols for hematopoietic differentiation, ESCs/iPS will readily give rise to differentiated hematopoietic cells as well as colony-forming cells.[93,94] However, HSCs capable of repopulating the hematopoietic system of lethally irradiated adult recipients are not generated. In this regard, differentiation recapitulates the development of the earliest embryonic hematopoietic tissue, the yolk sac.[95,96] Direct reprogramming provides the opportunity to induce human adult cells directly, possibly rendering homogenous populations. Inducing HSCs and other adult stem cells will potentially have the advantage to generate cells that self-renew, but since they are multipotent, their differentiation will be more straightforward. It remains an important question whether adult stem cells that retain self-renewal and multipotency properties are programmable.

To fully realize the potential of *in vitro* reprogrammed cells, we need to understand the molecular and epigenetic determinants that convert one cell type into another. Systematically defining the constellation of reprogramming factors provides the opportunity to allocate a master regulatory network for each cell type. The more the reprogramming factor combination resembles the target cell type regulatory network, the more efficient and accurate reprogramming would be. Taken together, these approaches will not only provide insights into development and disease but also a source of material to study and apply to both.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Simonsson, S. & J.B. Gurdon. 2005. Changing cell fate by nuclear reprogramming. *Cell Cycle* **4:** 513–515.
2. DiBerardino, M.A. 1988. Genomic multipotentiality of differentiated somatic cells. *Cell Differ. Dev*. **25**(Suppl): 129–136.
3. Surani, M.A. 2001. Reprogramming of genome function through epigenetic inheritance. *Nature* **414:** 122–128.
4. Yamanaka, S. & H.M. Blau. 2010. Nuclear reprogramming to a pluripotent state by three approaches. *Nature* **465:** 704–712.
5. Gurdon, J.B. 1962. Adult frogs derived from the nuclei of single somatic cells. *Dev. Biol*. **4:** 256–273.

6. Wilmut, I. *et al*. 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature* **385:** 810–813.
7. Hochedlinger, K. & R. Jaenisch. 2002. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* **415:** 1035–1038.
8. Inoue, K. *et al*. 2005. Generation of cloned mice by direct nuclear transfer from natural killer T cells. *Curr. Biol*. **15:** 1114–1118.
9. Harris, H. *et al*. 1969. Suppression of malignancy by cell fusion. *Nature* **223:** 363–368.
10. Blau, H.M. *et al*. 1985. Plasticity of the differentiated state. *Science* **230:** 758–766.
11. Chiu, C.P. & H.M. Blau. 1984. Reprogramming cell differentiation in the absence of DNA synthesis. *Cell* **37:** 879–887.
12. Tada, M. *et al*. 2001. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr. Biol*. **11:** 1553–1558.
13. Pereira, C.F. *et al*. 2008. Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2. *PLoS Genet*. **4:** e1000170.
14. Cowan, C.A. *et al*. 2005. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* **309:** 1369–1373.
15. Do, J.T. & H.R. Scholer. 2004. Nuclei of embryonic stem cells reprogram somatic cells. *Stem Cells* **22:** 941–949.
16. Byrne, J.A. *et al*. 2003. Nuclei of adult mammalian somatic cells are directly reprogrammed to oct-4 stem cell gene expression by amphibian oocytes. *Curr. Biol*. **13:** 1206–1213.
17. Wakayama, T. *et al*. 1998. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* **394:** 369–374.
18. Schneuwly, S., R. Klemenz & W.J. Gehring. 1987. Redesigning the body plan of Drosophila by ectopic expression of the homoeotic gene Antennapedia. *Nature* **325:** 816–818.
19. Davis, R.L., H. Weintraub & A.B. Lassar. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51:** 987–1000.
20. Takahashi, K. & S. Yamanaka. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126:** 663–676.
21. Aoi, T. *et al*. 2008. Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* 321**:** 699–702, doi: 10.1126/science.1154884.
22. Takahashi, K. *et al*. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131:** 861–872.
23. Feng, R. *et al*. 2008. PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells. *Proc. Natl. Acad. Sci. USA* **105:** 6057–6062.
24. Kajimura, S. *et al*. 2009. Initiation of myoblast to brown fat switch by a PRDM16-C/EBP-beta transcriptional complex. *Nature* **460:** 1154–1158.
25. Ieda, M. *et al*. 2010. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142:** 375–386.
26. Vierbuchen, T. *et al*. 2010. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463:** 1035–1041.
27. Huang, P. *et al*. 2011. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. *Nature* **475:** 386–389.
28. Zhou, Q. *et al*. 2008. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature* **455:** 627–632.
29. Anokye-Danso, F. *et al*. 2011. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* **8:** 376–388.
30. Ambasudhan, R. *et al*. 2011. Direct reprogramming of adult human fibroblasts to functional neurons under defined conditions. *Cell Stem Cell* **9:** 113–118.
31. Wernig, M. *et al*. 2007. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448:** 318–324.
32. Maherali, N. *et al*. 2007. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1:** 55–70.
33. Okita, K., T. Ichisaka & S. Yamanaka. 2007. Generation of germline-competent induced pluripotent stem cells. *Nature* **448:** 313–317.
34. Gurdon, J.B. & J.A. Byrne. 2003. The first half-century of nuclear transplantation. *Proc. Natl. Acad. Sci. USA* **100:** 8048–8052.
35. Stadtfeld, M., K. Brennand & K. Hochedlinger. 2008. Reprogramming of pancreatic Beta cells into induced pluripotent stem cells. *Curr. Biol*. **18:** 890–894.
36. Hanna, J. *et al*. 2008. Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* **133:** 250–264.
37. Kim, J.B. *et al*. 2008. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* **454:** 646–650.
38. Eminli, S. *et al*. 2008. Reprogramming of neural progenitor cells into iPS cells in the absence of exogenous Sox2 expression. *Stem Cells* **26:** 2467–2474.
39. Di Stefano, B., A. Prigione & V. Broccoli. 2008. Efficient genetic reprogramming of unmodified somatic neural progenitors uncovers the essential requirement of Oct4 and Klf4. *Stem Cells Dev.* **18:** 707–716.
40. Tsai, S.Y. *et al*. 2011. Single transcription factor reprogramming of hair follicle dermal papilla cells to induced pluripotent stem cells. *Stem Cells* **29:** 964–971.
41. Tsai, S.Y. *et al*. 2010. Oct4 and klf4 reprogram dermal papilla cells into induced pluripotent stem cells *Stem Cells*. **28:** 221–228.
42. Yu, J. *et al*. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318:** 1917–1920.
43. Park, I.H. *et al*. 2008. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451:** 141–146.
44. Lowry, W.E. *et al*. 2008. Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl. Acad. Sci. USA* **105:** 2883–2888.
45. Maherali, N. *et al*. 2008. A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* **3:** 340–345.
46. Sterneckert, J., S. Hoing & H.R. Scholer. 2011. Concise review: oct4 and more: the reprogramming expressway. *Stem Cells* **30:** 15–21.

47. Nakagawa, M. *et al*. 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol*. **26:** 101–106.
48. Wernig, M. *et al*. 2008. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* **2:** 10–12.
49. Feng, B. *et al*. 2009. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat. Cell Biol*. **11:** 197–203.
50. Guo, G. & A. Smith. 2010. A genome-wide screen in EpiSCs identifies Nr5a nuclear receptors as potent inducers of ground state pluripotency. *Development* **137:** 3185–3192.
51. Heng, J.C. *et al*. 2010. The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells *Cell Stem Cell* **6:** 167–174.
52. Singhal, N. *et al*. 2010. Chromatin-remodeling components of the BAF complex facilitate reprogramming. *Cell* **141:** 943–955.
53. Wang, T. *et al*. 2011. The histone demethylases jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. *Cell Stem Cell* **9:** 575–587.
54. Wang, W. *et al*. 2011. Rapid and efficient reprogramming of somatic cells to induced pluripotent stem cells by retinoic acid receptor gamma and liver receptor homolog 1. *Proc. Natl. Acad. Sci. USA* **108:** 18283–18288.
55. Han, J. *et al*. 2010. Tbx3 improves the germ-line competency of induced pluripotent stem cells. *Nature* **463:** 1096–1100.
56. Chia, N.Y. *et al*. 2010. A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* **468:** 316–320.
57. Maekawa, M. *et al*. 2011. Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* **474:** 225–229.
58. Zhao, Y. *et al*. 2008. Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell* **3:** 475–479.
59. Banito, A. *et al*. 2009. Senescence impairs successful reprogramming to pluripotent stem cells. *Genes Dev*. **23:** 2134–2139.
60. Mikkelsen, T.S. *et al*. 2008. Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454:** 49–55.
61. Pereira, C.F. *et al*. 2010. ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency. *Cell Stem Cell* **6:** 547–556.
62. Samavarchi-Tehrani, P. *et al*. 2010. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **7:** 64–77.
63. Li, R. *et al*. 2010. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* **7:** 51–63.
64. Egli, D., G. Birkhoff & K. Eggan. 2008. Mediators of reprogramming: transcription factors and transitions through mitosis. *Nat. Rev. Mol. Cell Biol*. **9:** 505–516.
65. Vierbuchen, T. & M. Wernig. 2011. Direct lineage conversions: unnatural but useful? *Nat. Biotechnol*. **29:** 892–907.
66. Faerman, A. *et al*. 1993. Ectopic expression of MyoD1 in mice causes prenatal lethalities. *Dev. Dyn*. **196:** 165–173.
67. Schafer, B.W. *et al*. 1990. Effect of cell history on response to helix-loop-helix family of myogenic regulators. *Nature* **344:** 454–458.
68. Xie, H. *et al*. 2004. Stepwise reprogramming of B cells into macrophages. *Cell* **117:** 663–676.
69. Pang, Z.P. *et al*. 2011. Induction of human neuronal cells by defined transcription factors. *Nature* **476:** 220–223.
70. Caiazzo, M. *et al*. 2011. Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* **476:** 224–227.
71. Pfisterer, U. *et al*. 2011. Direct conversion of human fibroblasts to dopaminergic neurons. *Proc. Natl. Acad. Sci. USA* **108:** 10343–10348.
72. Sekiya, S. & A. Suzuki. 2011. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* **475:** 390–393.
73. Szabo, E. *et al*. 2010. Direct conversion of human fibroblasts to multilineage blood progenitors. *Nature* **468:** 521–526.
74. Kim, J. *et al*. 2011. Direct reprogramming of mouse fibroblasts to neural progenitors. *Proc. Natl. Acad. Sci. USA* **108:** 7838–7843.
75. Efe, J.A. *et al*. 2011. Conversion of mouse fibroblasts into cardiomyocytes using a direct reprogramming strategy. *Nat. Cell Biol*. **13:** 215–222.
76. Ashburner, M. *et al*. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet*. **25:** 25–29.
77. Bhutani, N. *et al*. 2010. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463:** 1042–1047.
78. Vaquerizas, J.M. *et al*. 2009. A census of human transcription factors: function, expression and evolution. *Nat. Rev. Genet*. **10:** 252–263.
79. Hannenhalli, S. & K.H. Kaestner. 2009. The evolution of Fox genes and their role in development and disease. *Nat. Rev. Genet*. **10:** 233–240.
80. Jiang, J. *et al*. 2008. A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat. Cell Biol*. **10:** 353–360.
81. Rudnicki, M.A. *et al*. 1992. Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell* **71:** 383–390.
82. Rudnicki, M.A. *et al*. 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* **75:** 1351–1359.
83. Williams, R.R. *et al*. 2006. Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. *J. Cell Sci*. **119:** 132–140.
84. Zaret, K.S. *et al*. 2008. Pioneer factors, genetic competence, and inductive signaling: programming liver and pancreas progenitors from the endoderm. *Cold Spring Harb. Symp. Quant. Biol*. **73:** 119–126.
85. Zaret, K.S. & J.S. Carroll. 2011. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev*. **25:** 2227–2241.
86. Yan, J. *et al*. 2006. The forkhead transcription factor FoxI1 remains bound to condensed mitotic chromosomes and stably remodels chromatin structure. *Mol. Cell Biol*. **26:** 155–168.
87. Ghisletti, S. *et al*. 2010. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* **32:** 317–328.

88. Heinz, S. *et al*. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38:** 576–589.
89. Di Tullio, A. *et al*. 2011. CCAAT/enhancer binding protein alpha (C/EBP(alpha))-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation. *Proc. Natl. Acad. Sci. USA* **108:** 17016–17021.
90. Bussmann, L.H. *et al*. 2009. A robust and highly efficient immune cell reprogramming system. *Cell Stem Cell* **5:** 554–566.
91. Hanna, J. *et al*. 2009. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462:** 595–601.
92. Murry, C.E. & G. Keller. 2008. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132:** 661–680.
93. Doetschman, T.C. *et al*. 1985. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. Exp. Morphol*. **87:** 27–45.
94. Keller, G. *et al*. 1993. Hematopoietic commitment during embryonic stem cell differentiation in culture. *Mol. Cell Biol*. **13:** 473–486.
95. North, T. *et al*. 1999. Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development* **126:** 2563–2575.
96. Shalaby, F. *et al*. 1995. Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. *Nature* **376:** 62–66.
97. Weintraub, H. *et al*. 1989. Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc. Natl. Acad. Sci. USA* **86:** 5434–5438.