ANNUAL REVIEWS

Annual Review of Cell and Developmental Biology

Cell Reprogramming: The Many Roads to Success

Begüm Aydin[1,2] and Esteban O. Mazzoni[1,2]

[1]Department of Biology, New York University, New York, NY 10003, USA; email: eom204@nyu.edu

[2]Neuroscience Institute, Department of Neuroscience and Physiology, NYU School of Medicine, New York, NY 10016, USA

Annu. Rev. Cell Dev. Biol. 2019. 35:433–52

First published as a Review in Advance on July 23, 2019

The *Annual Review of Cell and Developmental Biology* is online at cellbio.annualreviews.org

https://doi.org/10.1146/annurev-cellbio-100818-125127

Keywords

cellular reprogramming strategies, transdifferentiation, mechanisms of direct reprogramming, differentiation trajectories, transcription factors

Abstract

Cellular reprogramming experiments from somatic cell types have demonstrated the plasticity of terminally differentiated cell states. Recent efforts in understanding the mechanisms of cellular reprogramming have begun to elucidate the differentiation trajectories along the reprogramming processes. In this review, we focus mainly on direct reprogramming strategies by transcription factors and highlight the variables that contribute to cell fate conversion outcomes. We review key studies that shed light on the cellular and molecular mechanisms by investigating differentiation trajectories and alternative cell states as well as transcription factor regulatory activities during cell fate reprogramming. Finally, we highlight a few concepts that we believe require attention, particularly when measuring the success of cell reprogramming experiments.

Contents

INTRODUCTION

During development, cells progressively acquire terminal fates by restricting their developmental potential. Although we usually consider the terminal states of development perdurable, cells can alter their fate by changing their developmental trajectory or by transdifferentiating into another terminally differentiated state without reverting to an early developmental stage (Jarriault et al. 2008, Küntziger & Collas 2004). Changes to cell fate specifications can also be linked to pathological cell states. For instance, several types of tumors dedifferentiate or acquire features of a cell type different from the cell types that originated the initial tumor mass (Friedmann-Morvinski & Verma 2014). Thus, terminal differentiation can be plastic, and the right set of extracellular and intracellular cues can drive cells out of their natural differentiation paths or states. Transdifferentiation or cell fate conversions have become a common experimental approach to test for the sufficiency of signaling molecules and transcription factors (TFs) to induce cellular phenotypes. As TF and signaling molecule cocktails became more complex over time, these gain-of-function experiments began acquiring the programming nomenclature. Thus, we define reprogramming as an unnatural differentiation trajectory that bypasses the progenitor stages. Reprogramming modifies the gene regulatory network of a resident cell type (a set of expressed genes and their interactions) without recapitulation of a natural developmental trajectory.

The visual clarity of Waddington's (1957) epigenetic landscape diagram resulted in its extensive use as a graphic representation of the development. Embryonic differentiation is represented by a ball (we will think of it as a single cell in this example) rolling downhill in a landscape. Representing the restriction of the differentiation potential, the ball's trajectory through the landscape is constrained, or canalized, by bifurcating valleys. Of note, Waddington (1957) acknowledged that the bifurcating paths are just for clarity and that pluripotent cells can be diverted to more than two differentiation paths. At the bottom of the landscape, the ball rests in one of many valleys representing the terminally differentiated cellular states. While the use of this representation for developmental studies is straightforward, there are a few concepts in Waddington's (1957) "The Cybernetics of Development" chapter (pp. 11–58) that are worth contemplating when we refer to reprogramming experiments with the Waddingtonian paradigm. First, development allows for controlled variations within the valleys. The robustness of the system maintaining a cell fate is represented by the cross sections of the epigenetic landscape. Thus, a steeper and narrower valley suggests a stable cell fate that would be more difficult to reprogram to another fate. Second, the shape of the epigenetic landscape is defined by the interactions among genes (**Figure 1**). Mutations do not affect the ball rolling down the hill but modify the shape of the landscape. By

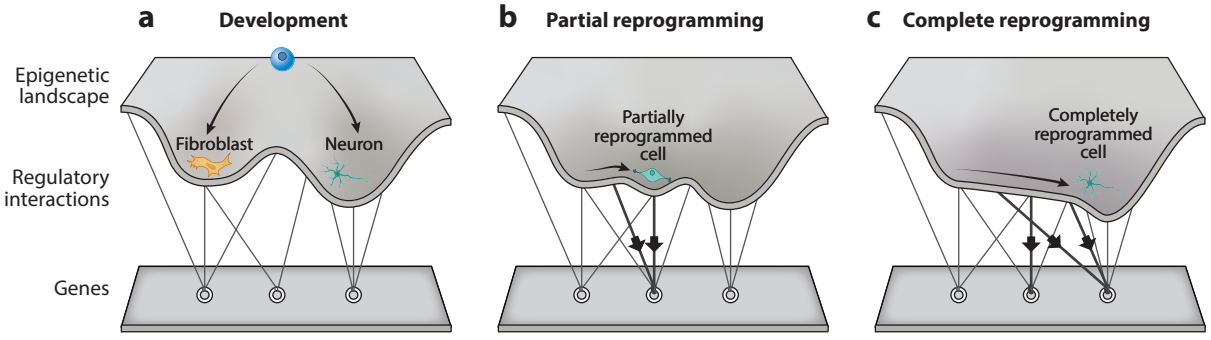

Figure 1

Waddington's (1957) epigenetic landscape applied to reprogramming experiments. (*a*) During normal development, the differentiation potential of a cell type is restricted by the bifurcating valleys that represent stable terminally differentiated cell states. These cell states are established by the complex gene regulatory interactions specific to each cell state. (*b*) Reprogramming (gain-of-function) experiments introduce novel gene regulatory interactions that were initially not present in the starting cell type and thus result in modification of the differentiation trajectories. Factors that are not potent enough to pull or push off the surfaces of the epigenetic landscape result in the generation of stable partially reprogrammed cell states. (*c*) Complete reprogramming of cell fate occurs when factors are potent enough to introduce novel gene regulatory interactions to facilitate the conversion of one cell fate to another by overcoming all barriers imposed by gene regulatory networks on the epigenetic landscape.

logical extrapolation, gain-of-function experiments aimed at reprogramming cell fates modify the landscape by generating novel valleys, or creodes as defined by Waddington (1957). This is a particularly important concept for reprogramming experiments. The modification of the landscape might create unnatural valleys for the ball to rest on. These novel stable states might represent the stable partially reprogrammed states often seen in reprogramming strategies (**Figure 1*b***). Third, developmental paths contain focal and peripheral factors in determination of the creode. Focal factors are required for the formation of novel valleys, since when they no longer operate, the tissue or organ does not develop normally, and the cells change their identity to something else (i.e., homeotic transformations in *Drosophila*). In contrast, peripheral factors cause the tissue or organ to develop abnormally but to remain recognizably similar to its normal form and are thus peripheral in the determination of the creode. Moreover, Waddington (1957) recognizes the virtue of positive feedback to induce substantial changes. Thus, focal genes that engage in positive feedback or reinforce a gene regulatory network are more likely to modify the epigenetic landscape and to induce transdifferentiation (**Figure 1*c***). As we discuss below, the milestone myogenic reprogramming experiments on MyoD1 fulfill this set of criteria.

While the reprogramming topic is rich and diverse, we discuss below a few outstanding issues in the field ranging from current methods to the expected outcomes of direct reprogramming. We mainly concentrate on the salient issues of direct reprogramming of mammalian cell types by accentuating the terminally differentiated cell states. The large volume of literature on neuronal reprogramming by different methods renders it a good case study for comparisons across strategies and mechanisms, and we thus use neuronal reprogramming as a primary example throughout this review. While not the focus of this review, plants and animals with high regenerative capacity are remarkable examples of developmental plasticity and reprogramming (Knapp & Tanaka 2012, Pierre-Jerome et al. 2018, Reddien 2018, Sugimoto et al. 2011).

DIRECT REPROGRAMMING STRATEGIES: EXTRACELLULAR CUES

Direct reprogramming relies on the application of an external force that disrupts the resident gene regulatory network and establishes a new one. Typically, such reprogramming can be done by the

activation or inhibition of signaling pathways or the forced expression of TFs. Direct reprogramming strategies find their origins in gain-of-function experiments aimed at testing sufficiency. From this point of view, reprogramming experiments test a specific hypothesis about the sufficiency of cellular environments, specific signals, or TFs to induce particular cell fates. Spemann & Mangold's (1923) transplantation experiments demonstrated that signals from a piece of the dorsal blastopore in the gastrula embryo are sufficient to induce cells on the ventral side to adopt fates that they normally would not. The organizer concept became a fundamental principle that governs our understanding of embryonic development, from digit patterning to generation of diverse neuronal cell types during embryogenesis, for example (Honig & Summerbell 1985, Placzek et al. 1990). These experiments also demonstrate that it is possible to derail normal development and to induce, or program, different progenitor cell types and ultimately different tissues by extracellular signals. Application of developmentally relevant signals, such as the ventralizing signal sonic hedgehog (SHH) and the anterior-posterior patterning signal retinoic acid (RA), can mimic development in vitro by differentiating mouse embryonic stem cells (ESCs) to spinal motor neurons at high efficiency (Wichterle et al. 2002). Applying directed differentiation by RA and SHH signaling to induced pluripotent stem cells (iPS cells) reprogrammed from fibroblasts allows for generation of patient-derived motor neurons (Dimos et al. 2008).

In principle, perturbation of signaling pathways can induce cellular reprogramming. Setting up a large screen to identify small-molecule cocktails that can induce cellular reprogramming requires large quantities of the initial cell type to be reprogrammed. Since mouse or human fibroblasts can easily be extracted and cultured in vitro, several protocols have been developed to differentiate fibroblasts into pluripotent or terminally differentiated cell types by small molecules. For example, treating fibroblasts with valproic acid (VPA), CHIR99021, and RepSox under hypoxic conditions reprograms them to neuronal progenitors (Cheng et al. 2014). Small-molecule cocktails with Forskolin, SP600125, GO6983, and Y-27632 or Forskolin, ISX9, CHIR99021, SB431542, and I-BET151 reprogram fibroblasts to postmitotic neurons instead (Hu et al. 2015, Li et al. 2015). There seems to be no unique neuronal chemical cocktail, since both chemical cocktails achieve the basic goal in generating cells that are postmitotic, that extend neurites, and that are electrically active.

We can extract a few important insights on reprogramming mechanisms from these neuronal reprogramming experiments by small molecules. Increasing cAMP levels (by Forskolin) and inhibiting GSK-3 (by CHIR99021) seem to be crucial in neuronal reprogramming since both chemical cocktails contain Forskolin and CHIR99021. Downregulation of the resident gene regulatory network is paramount for achieving complete cellular reprogramming. A cocktail of VPA (a histone deacetylase inhibitor), RepSox (a Tgf-β inhibitor), and CHIR99021 has also been used to facilitate the generation of pluripotent stem cells from fibroblasts (Huangfu et al. 2008, Ichida et al. 2009, Li et al. 2011). Thus, even though these chemicals are not specific to neuronal reprogramming, they may aid in downregulation of the resident gene regulatory network, and such downregulation allows for the establishment of the incoming network. In that vein, I-BET151 (a BET family bromodomain inhibitor) seems to be responsible for the downregulation of the fibroblast gene regulatory network; the molecular mechanism is not clear, but the process may occur through calcium signaling. Moreover, ISX9 is proposed to be responsible for inducing NeuroD1 and Neurog2, two strong inducers of neuronal fate (Li et al. 2015). Thus, the pleiotropic effect of small molecules seems well suited to perturbing the general cellular state, and such perturbation in turn allows specific signaling events or TFs to initiate cell fate conversion.

Reprogramming strategies that rely on small molecules do not require genetic modifications or viral vectors, and the molecules can be removed from the environment after successful reprogramming. Thus, reprogramming presents a promising strategy to generate cells for disease modeling,

drug discovery, and cell replacement. However, reprogramming by small molecules has challenges. While TFs control cell differentiation by binding to genomic DNA, small molecules or signaling factors touch on many biological processes. Signaling pathways are especially sensitive to cellular states, and signaling molecules may elicit different responses in different cell states. Additionally, as expected from developmentally regulated signaling pathways, signaling strength may require fine tuning. The powerful promise of the approach justifies the intensification of efforts aimed at producing precise and efficient reprogramming strategies by small molecules.

DIRECT REPROGRAMMING STRATEGIES: INTRACELLULAR CUES

Reprogramming by extracellular cues ultimately leads to the induction of intracellular forces, such as TFs and chromatin modifiers, that facilitate the conversion of cell fate. Thus, direct reprogramming strategies that exploit downstream intracellular processes are in principle easier to develop and, to date, have yielded more robust reprogramming outcomes across cell types. While the mechanisms by which extracellular reprogramming results in cellular conversion remain largely elusive, there is a wealth of information on TF-mediated direct reprogramming of cell types.

A notable example of reprogramming began with experiments pioneered by Briggs & King (1952), followed by the landmark experiments by Gurdon et al. (1958) in *Xenopus laevis.* These experiments demonstrated that transplantation of nuclei from cells at different stages into enucleated frog oocytes results in the generation of sexually mature frogs (Gurdon et al. 1958). Thus, a differentiated cell state is plastic, and the available cues in the oocyte cytoplasm can reprogram a differentiated cell nucleus to an undifferentiated or totipotent state. Reprogramming to pluripotency with intracellular cues highlights the gene regulatory plasticity of differentiated cells and is at the core of cloning experiments that culminated in the cloning of Dolly the sheep (Campbell et al. 1996).

The molecular identification of TFs as intracellular reprogramming cues began with the observation that hybrid cells obtained from ESC and fibroblast fusions acquire pluripotent phenotypes (Tada et al. 2001). Thus, similar to the cues in the oocyte cytoplasm, intracellular ESC cues can also reprogram a somatic nucleus to a pluripotent state. Decades later, Takahashi & Yamanaka (2006) demonstrated in their seminal work that reprogramming to pluripotency can be reduced to the activity of several TFs: Oct4, Klf4, Sox2, and c-Myc (OKSM or Yamanaka factors). The pluripotent stem cells obtained by reprogramming of somatic cell types were named iPS cells, and the TFs were referred to as reprogramming factors. Thus, terminal fate is far from an irreversible state, and relevant cues, including TFs, can induce dramatic cell fate transformations by reversing natural developmental trajectories.

While reprogramming to the pluripotent state is perhaps the most famous example of controlling cell fate by the forced expression of TFs, pluripotency is not the only state that is the subject of reprogramming strategies. A similar historical timeline can be drawn for reprogramming strategies that start and end in terminally differentiated states. In this context, we consider early chemical reprogramming experiments in fibroblast cultures. The addition of a small chemical, 5-azacytidine, to fibroblasts induced the expression of muscle-specific genes (Taylor & Jones 1979). In a clever experimental design, the source of this transformation was postulated to be a single gene encoding for the muscle-specific TF MyoD1 (Davis et al. 1987, Lassar et al. 1986). Indeed, it was demonstrated later that forced MyoD1 expression was sufficient to induce muscle differentiation markers in cells derived from all three germ layers (Davis et al. 1987, Weintraub et al. 1989). These landmark gain-of-function experiments demonstrated the power of TFs to reprogram terminally differentiated cell types.

As has been done for MyoD1 overexpression, countless developmental biology studies have tested for the sufficiency of a factor, or combinations of factors, in inducing cell fates. Arguably the most famous case is reprogramming to the pluripotent state by Yamanaka factors, but several direct reprogramming protocols have been developed to produce pancreatic β-cells, hepatocytes, cardiomyocytes, and even extraembryonic early trophoblast stem-like cells, among other cell types (Benchetrit et al. 2015, Qian et al. 2012, Sekiya & Suzuki 2011, Zhou et al. 2008). Forced expression of the proneural factors Ascl1 and Neurog2 can reprogram astroglial cells to neurons (Berninger et al. 2007). A screen of TFs expressed in neural lineages identified a TF combination of Ascl1, Brn2, and Myt1l as sufficient to reprogram mouse and human fibroblasts to neurons (Pang et al. 2011, Vierbuchen et al. 2010). These reprogrammed neurons express neuronal markers, adopt neuronal morphologies, generate action potentials, and form functional synapses. Ascl1 and Neurog2, by themselves or along with other neural lineage TFs, are powerful enough to reprogram several somatic cell types to neurons (Gascón et al. 2017; Heinrich et al. 2011, 2014; Karow et al. 2012, 2018).

TFs are not the only intracellular factors that can drive cellular reprogramming, and other regulatory signals such as microRNAs can reprogram cell fate. For example, expression of *miR-9/9** and *miR-124* in fibroblasts by themselves or with NeuroD2 induces neuron-specific gene expression (Yoo et al. 2011). Thus, the universe of possible genetically encoded factors keeps expanding, and reprogramming strategies that include TFs and other intrinsic regulatory factors have gained popularity in generating large numbers of clinically relevant cell types for possible future therapies. As the transgene delivering strategies evolve rapidly with nonintegrating viral vectors, reprogramming factors can be locally delivered into specific tissues, and their expression can be controlled by cell- or stage-specific regulatory elements to gain exquisite control over the reprogramming process.

DOES THE CELL OF ORIGIN MATTER FOR REPROGRAMMING SUCCESS?

The importance of the cell of origin in successful reprogramming was noted in the early stages of reprogramming research. It was observed in early amphibian cloning experiments that the ability of transplanted nuclei to promote normal development declines as development progresses. Thus, not all cells have equal potential to be reprogrammed by the enucleated oocyte. Similarly, not all cells respond equally to MyoD1 expression (Weintraub et al. 1989). Although MyoD1 overexpression induces muscle-specific genes when expressed in fibroblasts, it fails to strongly induce muscle markers in neuroblasts, B16 melanoma cells, HeLa cells, and a hepatoma cell line. Moreover, although rat neuroblasts and B16 melanoma cells gain expression of muscle markers, they retain features of the starting cell type after MyoD1 expression. Thus, since the beginning of reprogramming research, it has been clear that the cell of origin plays a paramount role in the success of reprogramming experiments (**Figure 2**). Depending on the target cell type to be reprogrammed, the distinct chromatin and cellular environment of the starting cell type can facilitate or impede the reprogramming process (Apostolou & Hochedlinger 2013, Iwafuchi-Doi & Zaret 2016).

Conceptually, the most straightforward reprogramming strategy is to start with a pluripotent or a multipotent cell type and to directly program a terminal cell fate by bypassing the intermediate stages that it would naturally undergo during development. These protocols often produce the most efficient cell fate conversions, as the chromatin and cellular contexts of the pluripotent cells are presumably more permissive for differentiation. However, even programming from different pluripotent states has its own challenges. Because of the differences in the gene regulatory networks governing cell fate specification, it is often difficult to translate the mouse reprogramming

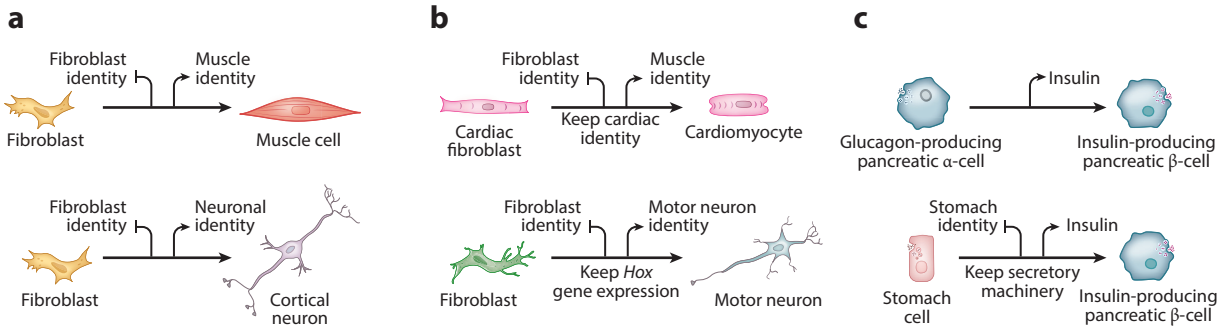

Figure 2

The ways to judge reprogramming success. (*a*) Examples of the ideal scenario of complete cell fate reprogramming, where the resident cell type is completely repressed and replaced by a new cellular identity. Complete replacement of gene expression with absolute precision of terminal fate is perhaps unachievable. (*b*) Partial cell reprogramming requires a fraction of the resident gene expression to be downregulated. Genes conferring the desirable cellular features can be maintained during the reprogramming process. Examples of this transformation include preserving the general tissue identity, for example, the cardiac identity (*top*), or preserving patterning genes, for example, *Hox* gene expression (*bottom*), during transdifferentiation. (*c*) Maintaining most cellular features and endowing these cells with a few desirable cell traits are possible. For example, acquiring the ability to produce insulin is sufficient for functional reprogramming (*top*). Reprogramming of one secretory cell type to another by maintaining the secretory machinery is beneficial to the reprogramming process.

strategies to reprogramming of human cells (Kriegstein et al. 2006, Lui et al. 2011, Schnerch et al. 2010). For example, while the Neurog2, Isl1, and Lhx3 TFs can robustly generate spinal motor neurons from mouse ESCs, the same combination of TFs are unable to induce mature motor neuron markers without extrinsic neuralizing signals when expressed in human ESCs (Hester et al. 2011, Mazzoni et al. 2013).

As the cell of origin becomes restricted in its developmental potential, the reprogramming process becomes challenging (**Figure 2*a***). Thus, reprogramming of somatic cells may require more than one TF, microRNA, or small molecule that enhances cell fate conversion (Ladewig et al. 2012, Liu et al. 2013, Vierbuchen et al. 2010, Yoo et al. 2011). For example, while Neurog2 expression in pluripotent cells is sufficient to induce neuronal fate, it fails to generate neurons when expressed in fibroblasts (Chanda et al. 2014, Liu et al. 2013, Meyer & Liu 2014). Ascl1 expression alone is sufficient to reprogram fibroblasts to neurons, but these neurons take much longer to acquire mature neuronal characteristics in the absence of Brn2 and Myt1l (Chanda et al. 2014, Vierbuchen et al. 2010). However, in this case, the addition of Brn2 and Myt1l seems to enhance neuronal maturation rather than aiding the neuronal conversion process. Similarly, reprogramming human fibroblasts to neurons is less efficient and requires the addition of NeuroD1 to the TF cocktail to induce neurons with mature characteristics (Pang et al. 2011, Son et al. 2011). The combination of the motor neuron TFs Neurog2, Isl1, and Lhx3 can induce spinal motor neurons with 98% conversion efficiency in mouse ESCs, but reprogramming of mouse fibroblasts to spinal motor neurons requires four additional TFs (Mazzoni et al. 2013, Son et al. 2011). Starting with cells that are developmentally related to the target cell type may result in a more feasible cell fate conversion. Related gene regulatory networks and chromatin landscapes would allow for the more efficient upregulation of genes associated with the new terminal fate. For example, astrocytes share a common progenitor with neurons and can thus be efficiently reprogrammed to neurons by forced expression of a single TF (Berninger et al. 2007, Chouchane et al. 2017, Gascón et al. 2017, Heinrich et al. 2010). Thus, since it is relatively easier to reprogram pluripotent cells or developmentally related differentiated cell types, these strategies can be advantageous in studying the molecular processes of diseases in vitro.

Reprogramming from neighboring tissues or resident cell types has the advantage of physical proximity to the injury site, favoring tissue integration and reprogramming in ideal conditions. For instance, reprogramming of human brain pericytes and resident astrocytes to neurons offers unique promise for cell replacement therapies in brain injury (Gascón et al. 2016; Heinrich et al. 2014; Karow et al. 2012; Pereira et al. 2017; Torper et al. 2013, 2015). In vivo reprogramming of cardiac fibroblasts to cardiomyocytes is a clear example of the advantages of in situ reprogramming (**Figure 2*b***) (Qian et al. 2012, Song et al. 2012). Cardiomyocytes reprogrammed from cardiac fibroblasts can rapidly integrate into tissues and improve cardiac function. Similar to the use of neural lineage–specific TFs in neuronal reprogramming, cardiomyocyte reprogramming protocols typically rely on the Gata4, Mef2c, and Tbx5 TFs, which control cardiac development (Bruneau 2013). Likewise, forced expression of the key TFs in pancreas development—Neurog3, Pdx1, and Mafa—reprograms pancreatic exocrine cells to insulin-secreting β-like cells in vivo (Zhou et al. 2008). Although with greater difficulty, terminally differentiated postmitotic neurons can be partially reprogrammed to different neuronal subtypes in vivo (Niu et al. 2018, Rouaux & Arlotta 2013). However, these partially reprogrammed neurons fail to downregulate the resident gene regulatory network. In general, neurons acquire a developmentally stable postmitotic fate that is maintained throughout the life of the organism. This stability may be reflected in their resistance toward being reprogrammed. Thus, the starting cell type constrains the differentiation trajectory as well as the effectiveness of the TF combinations in reprogramming cell fate. Not only the developmental distance of the cell of origin to the target cell but also the intrinsic cellular properties that can affect reprogramming efficiency should be considered. Partial downregulation of the resident gene regulatory program in the starting cell may be desirable in certain instances. The Neurog3, Pdx1, and Mafa TFs can also reprogram intestinal and stomach endocrine tissue to insulin-secreting β-like cells, and reprogrammed insulin-secreting β-like cells from stomach functionally reduce hyperglycemia (Ariyachet et al. 2016, Chen et al. 2014). A recent study showed that human islet non-β-cells (glucagon-producing α-cells and PPY-producing γ-cells) can be reprogrammed to insulin-producing β-cells by the Pdx1 and Mafa TFs (Furuyama et al. 2019). Reprogrammed β-cells secrete insulin upon glucose stimulation and reverse diabetes when transplanted into diabetic mice. Interestingly, most β-cells reprogrammed from α-cells acquired a hybrid bihormonal identity, retaining the α-cell gene expression profile. Thus, in this case, the reprogrammed cells maintain the resident endocrine regulatory network but acquire new secreted molecule and sensing mechanisms (**Figure 2*c***). Reprogramming of cells that have positional identity is another example in which partial downregulation of the resident gene regulatory program is advantageous (**Figure 2*b***). Fibroblasts and motor neurons are radically different cells in terms of their morphologies and physiological functions, yet both cell types have a strong positional identity that is imprinted in the *Hox* cluster chromatin. Thus, these cell types are not identical along the body axis (Philippidou & Dasen 2013, Rinn et al. 2007). Fibroblasts can be reprogrammed to motor neurons by a set of TFs, some of which are crucial in motor neuron development, and interestingly, motor neurons reprogrammed from caudal fibroblasts retain positional identity and express caudal *Hox* genes (Ichida et al. 2018, Son et al. 2011). In contrast, motor neurons directly programmed from pluripotent cells have all the core motor neuron genes, but not the *Hox* gene expression profile, which can be independently induced by signaling molecules (Mazzoni et al. 2013). This observation suggests that motor neuron fate can be decoupled into two separate dimensions: one that controls motor neuron fate transition and one that assigns cellular positional identity. While cells are not similar in their overall fate, they may have common features that are maintained during the reprogramming process. Thus, reprogramming outcome can be considered to be the combination of different gene regulatory networks that control specific aspects of cell physiology.

The relatively low efficiency of reprogramming protocols can be attributed to the reprogramming factors and the cell of origin. Reprogramming is not synchronous, even in a relatively homogeneous population. Thus, within one population, a fraction of cells is particularly receptive to reprogramming forces. The inefficiency of reprogramming to the pluripotent state led the research field to tackle the underlying mechanisms (Stadtfeld & Hochedlinger 2010, Yamanaka 2009). Live-imaging approaches were used in reprogramming of somatic cells to the pluripotent stem cell state by Yamanaka factors; the question was whether reprogramming to the pluripotent state happens in a stochastic or a determined manner. Reprogramming B cells or monocytes to iPS cells showed that most cells in the starting cell population have the potential to reprogram to pluripotency. Thus, stochastic changes facilitate cell fate conversion (Hanna et al. 2009). In contrast, differences in the subpopulations of starting cells can explain differences in reprogramming (Guo et al. 2014, Smith et al. 2010). An elegant study coupled cell tagging with single-cell RNA-seq to deconstruct the identity and the clonal history of cells undergoing direct reprogramming of mouse embryonic fibroblasts to induced endoderm progenitors (iEPs) by the Foxa1 and Hnf4α TFs (Biddy et al. 2018). Analysis of clonal history and cell identity at the single-cell resolution revealed that clones of the same lineage (cells that share ancestry) follow similar reprogramming trajectories. Thus, there seems to be no heterogeneity in the reprogramming success of clonally related cells. To tackle the question of heterogeneity of reprogramming processes, a recent study used single-cell RNA-seq sequencing to follow highly efficient cell fate conversion protocols applied to the same starting cell population: reprogramming of pre–B cells to macrophages (transdifferentiation) by the TF C/EBPα and reprogramming of pre–B cells to iPS cells by Yamanaka factors (Francesconi et al. 2019). The results revealed that heterogeneity in the reprogramming process arises in the starting cell population. Cells with low Myc activity efficiently transdifferentiate into macrophages but fail to reprogram to pluripotency, while cells with high Myc activity reprogram to pluripotency very efficiently but have much lower efficiency in transdifferentiating into macrophages. Thus, a heterogeneous starting cell population can be the source of noise during reprogramming experiments, thereby decreasing efficiency. However, such studies provide a platform to understand and manipulate factors controlling the reprogramming process.

Taken together, the evidence shows that the outcome of cellular reprogramming depends heavily on the starting cell type. Not only the right combination of factors but also the cellular context of the cell of origin result in successful reprogramming of cell fate. The starting cell population may be transcriptionally similar to the target cell type to be reprogrammed. The challenge for developmental biologists is to understand the different gene regulatory modules and ways to perturb each one independently. It seems tempting to conclude that developmentally or functionally related cells as the starting cell type can be a fruitful ground for reprogramming strategies. However, it is also necessary to consider the intrinsic plasticity of the starting cell type. Thus, taking advantage of distally related cells but plastic cell types with features similar to those of the target cell may be a useful strategy.

MECHANISMS OF CELLULAR REPROGRAMMING

While we know that the starting cell type plays a crucial role in the outcome of the reprogramming, the mechanisms at play have only begun to be elucidated. For complete cellular reprogramming to occur, the resident gene regulatory network should be partially or fully downregulated, and the gene regulatory network of the incoming target cell type should be upregulated. Reprogrammed cells may or may not follow a trajectory that resembles that of natural embryonic development. Understanding these cellular states along the reprogramming route is important because they have implications for disease modeling and regenerative medicine. While reprogrammed cells

that pass through mixed or proliferative cellular states could put patients receiving cell therapies at risk, such cellular states may allow unlimited expansion of cells to be used in directed differentiation strategies toward distinct cell fates. In addition to the lineage tracing experiments, single-cell RNA-seq technologies have been fruitful in discovering cellular differentiation trajectories during reprogramming in more detail.

Neuronal reprogramming of fibroblasts by small-molecule cocktails does not follow a canonical differentiation path. During normal development, mitotically active progenitor cells differentiate into postmitotic neuronal fates. However, during fibroblast-to-neuron reprogramming by small molecules, most cells do not incorporate BrdU, nor do they express canonical neuronal progenitor markers such as Pax6, Sox2, and Nestin (Hu et al. 2015, Li et al. 2015). These findings suggest that cell division or transitioning through neuronal progenitor–like states is not required for the acquisition of neuronal identity in chemically reprogrammed fibroblasts.

The ability to generate motor neurons from pluripotent stem cells by protocols that recapitulate their natural differentiation or by direct programming/reprogramming from ESCs and fibroblasts provides a unique platform to contrast reprogramming trajectories (Son et al. 2011, Velasco et al. 2017, Wichterle et al. 2002). Single-cell RNA-seq performed in direct programming of ESCs to motor neurons by forced expression of the spinal motor neuron TFs Neurog2, Isl1, and Lhx3 revealed a uniform differentiation trajectory (Velasco et al. 2017). Differentiating cells do not express Olig2, a typical motor neuron progenitor–stage gene marker, and thus bypass the embryonic progenitor stage (Mizuguchi et al. 2001, Novitch et al. 2001, Velasco et al. 2017). To understand the different differentiation trajectories of cells that transition through all differentiation intermediate stages versus those that are directly programmed to a terminal fate, Briggs et al. (2017) reconstructed and compared differentiation trajectories from single-cell RNA-seq time series experiments. Both directed differentiation (stepwise differentiation) and direct programming trajectories begin with ESCs and pass through a similar early neural progenitor state marked by the expression of Sox1 and Pax6. As differentiation proceeds, these two pathways diverge. Stepwise differentiating cells induce a genetic program that is associated with intermediate progenitor states, and this program becomes more restricted and mature over time. In contrast, directly programmed cells bypass the intermediate embryonic states and converge with the stepwise differentiation trajectory in the early motor neuron maturation state.

The upregulation of terminal regulatory programs in the early programming stages leads to rapid cell differentiation. The Ebf and Onecut TFs are postulated to be motor neuron terminal selector genes (Audouard et al. 2012, Francius & Clotman 2010, Kratsios et al. 2011, Roy et al. 2012). During motor neuron maturation, Onecut factors recruit Isl1 to enhancers that consolidate and maintain motor neuron gene expression (Rhee et al. 2016). Similarly, during direct motor neuron programming from ESCs, the Ebf and Onecut TFs are induced, and together they enable binding of spinal motor neuron TFs to previously inaccessible sites (Velasco et al. 2017). Thus, programming from pluripotent stem cells has analogous initial and final states while having intermediate states radically different from those of embryonic development. Cells subjected to direct programming from pluripotent states do not transition through either natural or unnatural states but experience a shortcut to differentiation by rapidly upregulating genes associated with a terminal differentiation state.

While it is conceptually easy to imagine how programming trajectories from pluripotent cells skip the intermediate developmental states, reprogramming from differentiated cell types—that is, transdifferentiation—may follow a differentiation trajectory different from that of embryonic development. Single-cell RNA-seq profiling of mouse embryonic fibroblasts reprogrammed to neurons revealed that reprogramming cells go through an intermediate progenitor state that differs from the canonical neural progenitor states observed in embryonic development and pluripotent

stem cell differentiation (Treutlein et al. 2016). Overlaying bulk neural stem cell transcriptomes onto the fibroblast-to-neuron single-cell RNA-seq differentiation trajectory shows that a fraction of cells that are in the intermediate position along the differentiation path express neural progenitor genes (such as *Nestin*, *Sox9*, and *Hes1*). Although cells in the intermediate states express neural progenitor markers, they do not express canonical neural progenitor marker genes such as *Sox2* and *Pax6* (Briggs et al. 2017). Thus, neurons reprogrammed from fibroblasts transition through an unnatural differentiation path with transient intermediate states.

A recent study analyzed the differentiation trajectory of direct pericyte-to-neuron reprogramming by Ascl1 and Sox2 and identified a set of genes whose expression dynamically changes during cell fate conversion (Karow et al. 2018). These switch genes are upregulated early during reprogramming, but their expression levels decline as neuronal reprogramming proceeds. These new intermediate steps are characterized by genes regulating developmental signaling pathways, such as *Notch2* and *Nog*, and are enriched in the germinal zones of the developing central nervous system, where the neural progenitor cells reside. Thus, reprogramming pericytes also pass through a transient neural progenitor–like state, and modulation of these cell signaling switch genes improves the efficiency of neuronal reprogramming. Similar to the case for fibroblast-to-neuron reprogramming, β-cells undergoing reprogramming do not transition through a rapidly dividing progenitor state or express genes associated with canonical progenitor states (Zhou et al. 2008). We cannot derive a unifying logic for differentiation trajectories from reprogramming experiments. However, we know it is possible to arrive at a terminal state without transitioning through canonical progenitor states. In most reprogramming experiments, strong downregulation of the resident gene regulatory network tends to occur in the early reprogramming states. Again, we highlight that the differentiation trajectory depends on the relationship between the cell of origin, the target cell type, and the reprogramming factors (e.g., signaling molecules, TFs of other genes).

ROADBLOCKS TO THE REPROGRAMMING PROCESS

Many factors impose roadblocks to or facilitate reprogramming. Some of these factors are shared by most reprogramming strategies; examples include mechanisms involved in cell cycle regulation, cellular senescence, genome stability, and repressive chromatin. Some are specific to a certain cell lineage, such as cell type–specific TFs that maintain the resident gene regulatory networks, factors that repress transition to specific fates (e.g., REST in non-neuronal cells), and metabolic regulators. These hurdles to reprogramming process were reviewed elsewhere (Gascón et al. 2017), and thus we focus here on a theme common to many reprogramming strategies.

En route to successful reprogramming, some cells may fail to fully downregulate the transcriptome of the starting cell type, resulting in a mixed or hybrid identity. Depending on the target cell type to be reprogrammed, retention of the starting cell type identity could be a desirable outcome (**Figure 2*b*,*c***). As described above, reprogramming of pancreatic non-β-cells to insulin-secreting β-cells is a great example; as long as the reprogrammed cells keep their secretory identity but start producing insulin, it does not matter whether they retain the ability to secrete glucagon (**Figure 2*c***) (Furuyama et al. 2019). However, retention of the starting cell type identity mostly suggests incomplete reprogramming and acquisition of a new stable cell state. Identifying roadblocks that impede the reprogramming process has been a long-time quest. Identifying ways to repress these roadblocks will greatly enhance differentiation. For example, NeuroD1 expression reprograms microglia to neurons (Matsuda et al. 2019). Its downstream targets, Scrt1 and Meis2, mediate repression of microglial gene expression through downregulation of key TFs in immune cell development and maintenance. Several computational approaches have been developed that aim to identify key reprogramming TFs and roadblocks to reliable

cell fate conversion (Cahan et al. 2014, D'Alessio et al. 2015, Okawa et al. 2016, Rackham et al. 2016). C/EBPα expression reprograms B cells to macrophage-like cells (Bussmann et al. 2009). In efforts to rapidly extinguish the resident gene regulatory network, the CellNet platform identified Pou2af1 and Ebf1 as two resident TFs that are maintained during reprogramming and that are responsible for preserving a fraction of B cell identity (Morris et al. 2014). Likewise, direct reprogramming of mouse fibroblasts to induced cardiomyocytes produced an intermediate cell state in which cells express both fibroblast and cardiomyocyte markers. In this case, the mRNA splicing factor *Ptbp1* was identified as the repressive barrier in the acquisition of cardiomyocyte fate, as deletion of this factor increased cardiomyocyte reprogramming efficiency (Liu et al. 2017). Tracking of reprogramming cells by CellTagging and single-cell RNA-seq during fibroblast-to-iEP conversion by the Foxa1 and Hnf4α TFs revealed two distinct trajectories: one that leads to successful reprogramming and another that leads to a dead-end state characterized by reexpression of fibroblast genes (Biddy et al. 2018). A putative methyltransferase, *Mettl7a1*, was significantly upregulated at the later stages along the successful reprogramming trajectory, and adding this gene to the reprogramming cocktail increased the yield of successfully reprogrammed endoderm progenitors.

During Ascl1-induced fibroblast-to-neuron reprogramming, Ascl1 also induces some neurons to adopt an aberrant myogenic fate (Treutlein et al. 2016). However, besides neuronal lineage–specific repressors such as REST and Groucho, the Myt1l TF canalizes neuronal differentiation during neuronal reprogramming (Jorgensen et al. 2009, Mall et al. 2017, Schoenherr & Anderson 1995). Hepatocyte-like cells (iHeps) can be reprogrammed by forced Hnf4α and FoxA expression in fibroblasts (Sekiya & Suzuki 2011). Analysis of gene regulatory networks induced by iHeps indicated that iHeps fail to extinguish the resident fibroblast program and weakly induce liver identity. In this case, the culprit is Cdx2, a reprogramming factor (Morris et al. 2014, Sekiya & Suzuki 2011). Thus, reprogramming can be blocked via many different mechanisms, making it hard to find a unifying strategy. Even some TFs that are required to reprogram cell fate produce undesirable side effects. Therefore, reprogramming strategies can be improved by selecting combinations of factors that are better in inducing a specific fate (such as focal factors, in Waddingtonian terms), in repressing the resident state, and in avoiding derailments during differentiation. If reprogramming factors induce additional unwanted features, addition of a TF that is not associated with the desired terminal fate but represses the aberrant features could be considered.

MOLECULAR MECHANISMS OF TRANSCRIPTION FACTOR ACTIVITY DURING REPROGRAMMING

Reprogramming of cell fate by TFs results in the rapid remodeling of the epigenetic and transcriptional landscape of the resident cell. We discuss here the mechanisms by which TFs engage with chromatin to initiate cell fate conversion, because some mechanistic roadblocks for TF-induced programming have been identified.

To activate the target cell gene regulatory network, TFs need to access developmentally silenced genes. Therefore, many reprogramming strategies use pioneer factors that can engage with the inaccessible chromatin and that can recruit chromatin remodelers and activators to upregulate transcriptional programs that lead to cell fate conversion (Iwafuchi-Doi & Zaret 2014, Zaret & Carroll 2011). Analysis of Oct4, Sox2, Klf4, and c-Myc genome-wide binding events in reprogramming of human fibroblasts to iPS cells showed that Oct4, Sox2, and Klf4 act as pioneer factors by binding to inaccessible chromatin regions (Soufi et al. 2012). c-Myc binding occurs in accessible chromatin regions and is not required for the reprogramming process, but c-Myc cooperatively enhances occupancy of Oct4, Sox2, and Klf4 sites where all factors are cobound. In-depth analysis

of Oct4 (a POU family TF), Sox2 (a HMG-box family TF), Klf4 (a zinc finger family TF), and c-Myc (a bHLH family TF) interactions with nucleosomes revealed that the pioneering activity of Oct4 and Sox2 relies on their ability to recognize a partial version of their canonical DNA binding motif on nucleosome-enriched inaccessible sites (Soufi et al. 2015). In this regard, Oct4 and Sox2 behave similarly to the well-characterized FoxA family of TFs, which can bind to nucleosomes via a DNA binding domain resembling that of a linker histone (Cirillo et al. 1998, 2002; Lee et al. 2005).

Master regulator bHLH TFs, such as MyoD and Ascl1, are strong inducers of cell fate reprogramming. Consistent with pioneering activity, Ascl1 directly associates with its target sequences in accessible and inaccessible regions during the reprogramming of fibroblasts to neurons (Wapinski et al. 2013). Ascl1's pioneering activity was attributed to its shorter basic helix 1 region, which was predicted to allow for its binding to nucleosomes by making loose contact (Soufi et al. 2015). Since Ascl1 has the shortest basic helix 1 region among the master regulator bHLH factors, it may have more potent pioneering activity. MyoD expression can initiate muscle differentiation in many cell types; however, MyoD was predicted to have weaker pioneering activity due to its longer basic helix 1, which would inhibit its binding efficiently to nucleosomes (Soufi et al. 2015). Nonetheless, MyoD can recruit histone acetyl transferases and chromatin remodelers (Tapscott 2005), and it associates with previously inaccessible sites (Casey et al. 2018). In contrast, the bHLH factor c-Myc, which contains a longer basic helix 1 domain, prefers to bind to accessible chromatin and does not seem to have pioneering activity (Soufi et al. 2015). Thus, even within a TF family whose members are commonly used to reprogram cell fate, the precise mode of action on chromatin is not uniform.

Both MyoD and NeuroD2 are in the bHLH family of TFs, yet they induce muscle and neuronal differentiation, respectively (Farah et al. 2000, Fong & Tapscott 2013, Vierbuchen et al. 2010). Thus, TFs with similar structural characteristics intrinsically encode the ability to initiate different cell fates. Analysis of the genome-wide binding events and transcriptional activity of these two bHLH factors showed that they bind to shared and private (specific) E-box sequence variants (Fong et al. 2012). MyoD and NeuroD2 are 39% identical in the amino acid sequence of their DNA binding domains (the bHLH domain). A chimeric protein in which the bHLH domain of MyoD is swapped with that of NeuroD2 converts MyoD to a neurogenic factor by binding to NeuroD2 private sites and activating the muscle-specific gene expression program in P19 carcinoma cells (Fong et al. 2015). With these examples, we highlight the need to understand TF regulatory activity at a level that can predict the intrinsic activity of each TF alone, and in combination, to improve reprogramming strategies.

Cell fate is specified by the combinatorial activity of TFs. Thus, reprogramming strategies typically rely on the simultaneous expression of a handful of TFs. In vertebrates, Neurog2 and Ascl1 induce differentiation of the vast majority of neurons during development and are thus widely used to induce neuronal fate (Bertrand et al. 2002, Guillemot & Hassan 2017). The prevailing model suggests that the postmitotic generic, or pan-neuronal, neuronal fate is patterned into neuronal subtypes by the activity of terminal selectors or TF combinations (Bhati et al. 2008; Doitsidou et al. 2013; Flames & Hobert 2011; Hobert 2008, 2011; Stefanakis et al. 2015). For example, during spinal motor neuron differentiation, Neurog2 activates pan-neuronal genes, while Isl1 and Lhx3 induce the motor neuron–specific gene regulatory program (Lee & Pfaff 2003, Lee et al. 2012, Mazzoni et al. 2013, Mizuguchi et al. 2001). In contrast, Ascl1 with Lmx1a and Nurr1 (or Foxa1) reprogram fibroblasts to a dopaminergic neuronal fate (Caiazzo et al. 2011, Pfisterer et al. 2011). While much attention has been paid to the Isl1-Lhx3 versus Lmx1a-Nurr1 (or Lmx1a-Foxa1) combinations in canalizing subtype-specific fate, Ascl1 and Neurog2 also contribute to neuronal subtype identities (Parras et al. 2002). A recent large-scale

unbiased screen tested 598 pairs of TFs, specifically the abilities of these TF pairs to induce neuronal identity in mouse fibroblasts (Tsunemoto et al. 2018). This study found that more than 12% (76 out of 598) of the TF pairs can reprogram fibroblasts to mature neurons with electrical activity and synaptic connectivity. Interestingly, gene expression analysis revealed that, even when other TFs are expressed, neuronal subtype identity is dominated by the Acheate-Scute (Ascl) or the Neurogenin (Neurog) proneural factor family. Thus, the activities of Ascl1 and Neurog2 in controlling neuronal subtype identities are magnified when these TFs are expressed in an ectopic environment in reprogramming protocols. Their contribution to neuronal subtype identities can be explained by the divergent binding patterns of Ascl1 and Neurog2, which in turn establishes distinct chromatin landscapes that affect the activity of downstream TFs in neuronal subtype specification (Aydin et al. 2019, Wapinski et al. 2013). In summary, cooperativity among neuronal reprogramming TFs can be indirectly achieved by their activity in the chromatin landscape.

CONCLUDING REMARKS

In addition to having a role in regenerative medicine by allowing for the generation of an unlimited number of cells for disease modeling, direct reprogramming strategies have become an advantageous platform to investigate specific functions of signaling molecules, TFs, and other factors that shape gene regulatory networks. Although much attention has been given to the positive regulators of cell fate, successful reprogramming also relies on repressing the resident cell fate. Thus, reprogramming strategies must overcome the mechanisms that actively prevent cells from weakening their resident cell fate. Chemical reprogramming by small molecules, in this sense, has taken advantage of the compounds that disrupt the resident gene regulatory networks by predominantly targeting chromatin modifiers. In contrast, forced TF expression seems to be a potent and precise reprogramming strategy. We believe that combining chemicals and TFs, or other genetically encoded signals, could strengthen reprogramming of terminal cell fates as it has strengthened reprogramming of pluripotency.

In the ideal reprogramming process, the cell of origin and the right reprogramming strategy trigger an unnatural differentiation path that culminates in successful cell fate conversion. With regard to mechanisms of various reprogramming strategies, finding a unifying model is difficult. The relationship between the cell of origin, the target cell type, and the reprogramming method introduces enough variation that examples of cells transitioning through different states can be found throughout the literature (Reid & Tursun 2018). Reprogramming strategies that start with either (*a*) a cell type similar to the target cell type or (*b*) a cell type that already expresses the desirable genes may take advantage of these partially reprogrammed states to maintain the desirable gene regulatory networks. However, we believe that the relationship between cells should be considered beyond lineage or superficial cellular similarities (**Figure 2**). Rather than selecting cells by lineage relationships, paying close attention to desirable cellular features in the cell of origin might increase the success rate of reprogramming strategies for clinical applications.

We believe that the field should first openly recognize the fact that generating an identical cellular copy of cells produced during embryonic development is in most cases not necessary, and perhaps unachievable. Unlike the case for many other cellular fates, reprogramming to pluripotency has a clear benchmark, and thus it is often used as a guiding example of cellular reprogramming. If successfully reprogrammed, iPS cells can give rise to all cells in the developing embryo. But perhaps a more modest set of criteria could be applied to reprogramming of other cell types. Few reviews have highlighted the minimal requirements for a cell to be considered successfully reprogrammed to a neuron (Drouin-Ouellet et al. 2017, Yang et al. 2011). Many of these requirements for successful neuronal differentiation (such as expression of action potential firing) may

suffice when these reprogrammed neurons are used to study basic neuronal physiology. However, a successful neuronal maturation status is necessary if the object of study is aging or late-onset neurodegenerative diseases. Similarly, a cell that induces a fraction of genes expressed in a pancreatic β-cell and that can regulate glycemia without safety concerns could be considered a success for cell replacement therapies, but not for the study of β-cell physiology. Therefore, the success of a reprogramming strategy should be defined by the goal in a given scenario.

Cellular reprogramming, by definition, relies on signaling molecules and TF operating in an ectopic cellular state as opposed to the states that they encounter during development. There is a profound gap between our understanding of developmental biology and the number of cell fates that can be efficiently reprogrammed due to our lack of precise understanding of how the developmental rules translate into reprogramming strategies. We believe that the field has collected a significant amount of data that can be mined to address this question. There is a large collection of unpublished failed reprogramming experiments. The novel cellular states produced in failed reprogramming experiments are likely to contain both conserved and novel gene regulatory programs that could be instrumental in understanding the compartmentalization of cell fate within regulatory networks. As in many fields, a repository of gene expression profiles of failed reprogramming experiments could be mined to reconstruct gene regulatory circuits.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We apologize to those whose work may not have been cited because of length constraints or our own ignorance. We would like to thank Ramin Rahni for the figures. Work in the authors' lab is supported by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (R01HD079682). B.A. has been funded by a NYSTEM predoctoral training grant (C026880) and the MacCracken and Dean's Dissertation Fellowship from NYU.

LITERATURE CITED

Apostolou E, Hochedlinger K. 2013. Chromatin dynamics during cellular reprogramming. *Nature* 502(7472):462–71

Ariyachet C, Tovaglieri A, Xiang G, Lu J, Shah MS, et al. 2016. Reprogrammed stomach tissue as a renewable source of functional β cells for blood glucose regulation. *Cell Stem Cell* 18(3):410–21

Audouard E, Schakman O, René F, Huettl RE, Huber AB, et al. 2012. The Onecut transcription factor HNF-6 regulates in motor neurons the formation of the neuromuscular junctions. *PLOS ONE* 7(12):e50509

Aydin B, Kakumanu A, Rossillo M, Moreno-Estelles M, Garipler G, et al. 2019. Proneural factors Ascl1 and Neurog2 contribute to neuronal subtype identities by establishing distinct chromatin landscapes. *Nat. Neurosci.* 22:897–908

Benchetrit H, Herman S, Van Wietmarschen N, Wu T, Makedonski K, et al. 2015. Extensive nuclear reprogramming underlies lineage conversion into functional trophoblast stem-like cells. *Cell Stem Cell* 17(5):543–56

Berninger B, Costa MR, Koch U, Schroeder T, Sutor B, et al. 2007. Functional properties of neurons derived from in vitro reprogrammed postnatal astroglia. *J. Neurosci.* 27(32):8654–64

Bertrand N, Castro DS, Guillemot F. 2002. Proneural genes and the specification of neural cell types. *Nat. Rev. Neurosci.* 3(7):517–30

Bhati M, Lee C, Nancarrow AL, Lee M, Craig VJ, et al. 2008. Implementing the LIM code: the structural basis for cell type–specific assembly of LIM-homeodomain complexes. *EMBO J.* 27(14):2018–29

Biddy BA, Kong W, Kamimoto K, Guo C, Waye SE, et al. 2018. Single-cell mapping of lineage and identity in direct reprogramming. *Nature* 564:219–24

Briggs J, Li V, Lee S, Woolf C, Klein A, Kirschner M. 2017. Mouse embryonic stem cells can differentiate via multiple paths to the same state. *eLife* 6:e26945

Briggs R, King TJ. 1952. Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. *PNAS* 38(5):455–63

Bruneau BG. 2013. Signaling and transcriptional networks in heart development and regeneration. *Cold Spring Harb. Perspect. Biol.* 5:a008292

Bussmann LH, Schubert A, Vu Manh TP, De Andres L, Desbordes SC, et al. 2009. A robust and highly efficient immune cell reprogramming system. *Cell Stem Cell* 5(5):554–66

Cahan P, Li H, Morris SA, Lummertz Da Rocha E, Daley GQ, Collins JJ. 2014. CellNet: network biology applied to stem cell engineering. *Cell* 158(4):903–15

Caiazzo M, Maria D, Dvoretskova E, Lazarevic D, Taverna S, et al. 2011. Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* 476(7359):224–27

Campbell KHS, McWhir J, Wilmut RI. 1996. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* 380:64–66

Casey BH, Kollipara RK, Pozo K, Johnson JE. 2018. Intrinsic DNA binding properties demonstrated for lineage-specifying basic helix-loop-helix transcription factors. *Genome Res.* 28:484–96

Chanda S, Ang CE, Davila J, Pak C, Mall M, et al. 2014. Generation of induced neuronal cells by the single reprogramming factor ASCL1. *Stem Cell Rep.* 3(2):282–96

Chen YJ, Finkbeiner SR, Weinblatt D, Emmett MJ, Tameire F, et al. 2014. De novo formation of insulin-producing "neo-β cell islets" from intestinal crypts. *Cell Rep.* 6(6):1046–58

Cheng L, Hu W, Qiu B, Zhao J, Yu Y, et al. 2014. Generation of neural progenitor cells by chemical cocktails and hypoxia. *Cell Res.* 24(6):665–79

Chouchane M, Melo de Farias AR, de Sousa Moura DM, Hilscher MM, Schroeder T, et al. 2017. Lineage reprogramming of astroglial cells from different origins into distinct neuronal subtypes. *Stem Cell Rep.* 9(1):162–76

Cirillo L, Lin F, Cuesta I, Friedman D, Jarnik M, Zaret KS. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* 9(2):279–89

Cirillo LA, McPherson CE, Bossard P, Stevens K, Cherian S, et al. 1998. Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO J.* 17(1):244–54

D'Alessio AC, Fan ZP, Wert KJ, Baranov P, Cohen MA, et al. 2015. A systematic approach to identify candidate transcription factors that control cell identity. *Stem Cell Rep.* 5(5):763–75

Davis RL, Weintraub H, Lassar AB. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51(6):987–1000

Dimos JT, Rodolfa KT, Niakan KK, Weisenthal LM, Mitsumoto H, et al. 2008. Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* 321(5893):1218–21

Doitsidou M, Flames N, Topalidou I, Abe N, Felton T, et al. 2013. A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in *C. elegans*. *Genes Dev.* 27(12):1391–405

Drouin-Ouellet J, Pircs K, Barker RA, Jakobsson J, Parmar M. 2017. Direct neuronal reprogramming for disease modeling studies using patient-derived neurons: What have we learned? *Front. Neurosci.* 11:530

Farah MH, Olson J, Sucic H, Hume R, Tapscott S, Turner D. 2000. Generation of neurons by transient expression of neural bHLH proteins in mammalian cells. *Development* 127:693–702

Flames N, Hobert O. 2011. Transcriptional control of the terminal fate of monoaminergic neurons. *Annu. Rev. Neurosci.* 34:153–84

Fong AP, Tapscott SJ. 2013. Skeletal muscle programming and re-programming. *Curr. Opin. Genet. Dev.* 23(5):568–73

Fong AP, Yao Z, Zhong JW, Cao Y, Ruzzo WL, et al. 2012. Genetic and epigenetic determinants of neurogenesis and myogenesis. *Dev. Cell* 22(4):721–35

Fong AP, Yao Z, Zhong JW, Johnson NM, Farr GH, et al. 2015. Conversion of MyoD to a neurogenic factor: Binding site specificity determines lineage. *Cell Rep*. 10(12):1937–46

Francesconi M, Di Stefano B, Berenguer C, de Andrés-Aguayo L, Plana-Carmona M, et al. 2019. Single cell RNA-seq identifies the origins of heterogeneity in efficient cell transdifferentiation and reprogramming. *eLife* 8:e41627

Francius C, Clotman F. 2010. Dynamic expression of the Onecut transcription factors HNF-6, OC-2 and OC-3 during spinal motor neuron development. *Neuroscience* 165(1):116–29

Friedmann-Morvinski D, Verma IM. 2014. Dedifferentiation and reprogramming: origins of cancer stem cells. *EMBO Rep*. 15(3):244–53

Furuyama K, Chera S, van Gurp L, Oropeza D, Ghila L, et al. 2019. Diabetes relief in mice by glucose-sensing insulin-secreting human α-cells. *Nature* 567(7746):43–48

Gascón S, Masserdotti G, Russo GL, Götz M. 2017. Direct neuronal reprogramming: achievements, hurdles, and new roads to success. *Cell Stem Cell* 21(1):18–34

Gascón S, Murenu E, Masserdotti G, Ortega F, Russo GL, et al. 2016. Identification and successful negotiation of a metabolic checkpoint in direct neuronal reprogramming. *Cell Stem Cell* 18(3):396–409

Guillemot F, Hassan BA. 2017. Beyond proneural: emerging functions and regulations of proneural proteins. *Curr. Opin. Neurobiol.* 42:93–101

Guo S, Zi X, Schulz VP, Cheng J, Zhong M, et al. 2014. Nonstochastic reprogramming from a privileged somatic cell state. *Cell* 156(4):649–62

Gurdon JB, Elsdale TR, Fischberg M. 1958. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* 182(4627):64–65

Hanna J, Saha K, Pando B, van Zon J, Lengner CJ, et al. 2009. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* 462(7273):595–601

Heinrich C, Bergami M, Gasco S, Dimou L, Sutor B, et al. 2014. *Sox2*-mediated conversion of NG2 glia into induced neurons in the injured adult cerebral cortex. *Stem Cell Rep*. 3:1000–14

Heinrich C, Blum R, Gascón S, Masserdotti G, Tripathi P, et al. 2010. Directing astroglia from the cerebral cortex into subtype specific functional neurons. *PLOS Biol.* 8(5):e1000373

Heinrich C, Gascón S, Masserdotti G, Lepier A, Sanchez R, et al. 2011. Generation of subtype-specific neurons from postnatal astroglia of the mouse cerebral cortex. *Nat. Protoc.* 6(2):214–28

Hester ME, Murtha MJ, Song S, Rao M, Miranda CJ, et al. 2011. Rapid and efficient generation of functional motor neurons from human pluripotent stem cells using gene delivered transcription factor codes. *Mol. Ther.* 19(10):1905–12

Hobert O. 2008. Regulatory logic of neuronal diversity: terminal selector genes and selector motifs. *PNAS* 105(51):20067–71

Hobert O. 2011. Regulation of terminal differentiation programs in the nervous system. *Annu. Rev. Cell Dev. Biol.* 27:681–96

Honig LS, Summerbell D. 1985. Maps of strength of positional signalling activity in the developing chick wing bud. *Development* 87:163–74

Hu W, Qiu B, Guan W, Wang Q, Wang M, et al. 2015. Direct conversion of normal and Alzheimer's disease human fibroblasts into neuronal cells by small molecules. *Cell Stem Cell* 17(2):204–12

Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, et al. 2008. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* 26(7):795–97

Ichida JK, Blanchard J, Lam K, Son EY, Chung JE, et al. 2009. A small-molecule inhibitor of Tgf-β signaling replaces *Sox2* in reprogramming by inducing *Nanog*. *Cell Stem Cell* 5(5):491–503

Ichida JK, Staats KA, Davis-Dusenbery BN, Clement K, Galloway KE, et al. 2018. Comparative genomic analysis of embryonic, lineage-converted, and stem cell-derived motor neurons. *Development* 9:dev168617

Iwafuchi-Doi M, Zaret KS. 2014. Pioneer transcription factors in cell reprogramming. *Genes Dev*. 28(24):2679–92

Iwafuchi-Doi M, Zaret KS. 2016. Cell fate control by pioneer transcription factors. *Development* 143(11):1833–37

Jarriault S, Schwab Y, Greenwald I. 2008. A *Caenorhabditis elegans* model for epithelial-neuronal transdifferentiation. *PNAS* 105(10):3790–95

Jorgensen HF, Terry A, Beretta C, Pereira CF, Leleu M, et al. 2009. REST selectively represses a subset of RE1-containing neuronal genes in mouse embryonic stem cells. *Development* 136(5):715–21

Karow M, Camp JG, Falk S, Gerber T, Pataskar A, et al. 2018. Direct pericyte-to-neuron reprogramming via unfolding of a neural stem cell-like program. *Nat. Neurosci.* 21:932–40

Karow M, Sanchez R, Schichor C, Masserdotti G, Ortega F, et al. 2012. Reprogramming of pericyte-derived cells of the adult human brain into induced neuronal cells. *Cell Stem Cell* 11:471–76

Knapp D, Tanaka EM. 2012. Regeneration and reprogramming. *Curr. Opin. Genet. Dev.* 22(5):485–93

Kratsios P, Stolfi A, Levine M, Hobert O. 2011. Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. *Nat. Neurosci.* 15:205–14

Kriegstein A, Noctor S, Martínez-Cerdeño V. 2006. Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nat. Rev. Neurosci.* 7:883–90

Küntziger T, Collas P. 2004. Transdifferentiation. In *Handbook of Stem Cells*, ed. R Lanza, H Blau, J Gearhart, B Hogan, D Melton, et al., pp. 147–51. Burlington, MA: Elsevier Acad. Press

Ladewig J, Mertens J, Kesavan J, Doerr J, Poppe D, et al. 2012. Small molecules enable highly efficient neuronal conversion of human fibroblasts. *Nat. Methods* 9(6):575–78

Lassar AB, Paterson BM, Weintraub H. 1986. Transfection of a DNA locus that mediates the conversion of 10T12 fibroblasts to myoblasts. *Cell* 47(5):649–56

Lee CS, Friedman JR, Fulmer JT, Kaestner KH. 2005. The initiation of liver development is dependent on Foxa transcription factors. *Nature* 435(7044):944–47

Lee S, Cuvillier JM, Lee B, Shen R, Lee JW, Lee S-K. 2012. Fusion protein Isl1-Lhx3 specifies motor neuron fate by inducing motor neuron genes and concomitantly suppressing the interneuron programs. *PNAS* 109(9):3383–88

Lee SK, Pfaff SL. 2003. Synchronization of neurogenesis and motor neuron specification by direct coupling of bHLH and homeodomain transcription factors. *Neuron* 38(5):731–45

Li X, Zuo X, Jing J, Ma Y, Wang J, et al. 2015. Small-molecule-driven direct reprogramming of mouse fibroblasts into functional neurons. *Cell Stem Cell* 17(2):195–203

Li Y, Zhang Q, Yin X, Yang W, Du Y, et al. 2011. Generation of iPSCs from mouse fibroblasts with a single gene, *Oct4*, and small molecules. *Cell Res.* 21(1):196–204

Liu M-LL, Zang T, Zou Y, Chang JC, Gibson JR, et al. 2013. Small molecules enable neurogenin 2 to efficiently convert human fibroblasts into cholinergic neurons. *Nat. Commun.* 4:2183

Liu Z, Wang L, Welch JD, Ma H, Zhou Y, et al. 2017. Single-cell transcriptomics reconstructs fate conversion from fibroblast to cardiomyocyte. *Nature* 551(7678):100–4

Lui JH, Hansen DV, Kriegstein AR. 2011. Development and evolution of the human neocortex. *Cell* 146(1):18–36

Mall M, Kareta MS, Chanda S, Ahlenius H, Perotti N, et al. 2017. Myt1l safeguards neuronal identity by actively repressing many non-neuronal fates. *Nature* 544:245–49

Matsuda T, Irie T, Katsurabayashi S, Hayashi Y, Nagai T, et al. 2019. Pioneer factor NeuroD1 rearranges transcriptional and epigenetic profiles to execute microglia-neuron conversion. *Neuron* 101(3):472–85

Mazzoni EO, Mahony S, Closser M, Morrison CA, Nedelec S, et al. 2013. Synergistic binding of transcription factors to cell-specific enhancers programs motor neuron identity. *Nat. Neurosci.* 16(9):1219–27

Meyer CA, Liu XS. 2014. Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. *Nat. Rev. Genet.* 15(11):709–21

Mizuguchi R, Sugimori M, Takebayashi H, Kosako H, Nagao M, et al. 2001. Combinatorial roles of Olig2 and Neurogenin2 in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. *Neuron* 31(5):757–71

Morris SA, Cahan P, Li H, Zhao AM, San Roman AK, et al. 2014. Dissecting engineered cell types and enhancing cell fate conversion via CellNet. *Cell* 158(4):889–902

Niu W, Zang T, Wang LL, Zou Y, Zhang CL. 2018. Phenotypic reprogramming of striatal neurons into dopaminergic neuron-like cells in the adult mouse brain. *Stem Cell Rep.* 11(5):1156–70

Novitch BG, Chen AI, Jessell TM. 2001. Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. *Neuron* 31:773–89

Okawa S, Nicklas S, Zickenrott S, Schwamborn JC, del Sol A. 2016. A generalized gene-regulatory network model of stem cell differentiation for predicting lineage specifiers. *Stem Cell Rep*. 7(3):307–15

Pang ZP, Yang N, Vierbuchen T, Ostermeier A, Fuentes DR, et al. 2011. Induction of human neuronal cells by defined transcription factors. *Nature* 476:220–23

Parras CM, Schuurmans C, Scardigli R, Kim J, Anderson DJ, Guillemot F. 2002. Divergent functions of the proneural genes *Mash1* and *Ngn2* in the specification of neuronal subtype identity. *Genes Dev*. 16(3):324–38

Pereira M, Birtele M, Shrigley S, Benitez J, Hedlund E, et al. 2017. Direct reprogramming of resident NG2 glia into neurons with properties of fast-spiking parvalbumin-containing interneurons. *Stem Cell Rep*. 9(3):742–51

Pfisterer U, Kirkeby A, Torper O, Wood J, Nelander J, et al. 2011. Direct conversion of human fibroblasts to dopaminergic neurons. *PNAS* 108(25):10343–48

Philippidou P, Dasen JS. 2013. *Hox* genes: choreographers in neural development, architects of circuit organization. *Neuron* 80(1):12–34

Pierre-Jerome E, Drapek C, Benfey PN. 2018. Regulation of division and differentiation of plant stem cells. *Annu. Rev. Cell Dev. Biol.* 34(1):289–310

Placzek M, Tessier-Lavigne M, Yamada T, Jessell T, Dodd J. 1990. Mesodermal control of neural cell identity: floor plate induction by the notochord. *Science* 250:985–88

Qian L, Huang Y, Spencer CI, Foley A, Vedantham V, et al. 2012. In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. *Nature* 485:593–98

Rackham OJL, Firas J, Fang H, Oates ME, Holmes ML, et al. 2016. A predictive computational framework for direct reprogramming between human cell types. *Nat. Genet.* 48(3):331–35

Reddien PW. 2018. The cellular and molecular basis for planarian regeneration. *Cell* 175(2):327–45

Reid A, Tursun B. 2018. Transdifferentiation: Do transition states lie on the path of development? *Curr. Opin. Syst. Biol.* 11:18–23

Rhee HS, Closser M, Guo Y, Bashkirova EV, Tan GC, et al. 2016. Expression of terminal effector genes in mammalian neurons is maintained by a dynamic relay of transient enhancers. *Neuron* 92(6):1252–65

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, et al. 2007. Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs. *Cell* 129(7):1311–23

Rouaux C, Arlotta P. 2013. Direct lineage reprogramming of post-mitotic callosal neurons into corticofugal neurons in vivo. *Nat. Cell Biol.* 15(2):214–21

Roy A, Francius C, Rousso DL, Seuntjens E, Debruyn J, et al. 2012. Onecut transcription factors act upstream of Isl1 to regulate spinal motoneuron diversification. *Development* 139(17):3109–19

Schnerch A, Cerdan C, Bhatia M. 2010. Distinguishing between mouse and human pluripotent stem cell regulation: the best laid plans of mice and men. *Stem Cells* 28:419–30

Schoenherr CJ, Anderson DJ. 1995. The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science* 267(5202):1360–63

Sekiya S, Suzuki A. 2011. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* 475(7356):390–95

Smith ZD, Nachman I, Regev A, Meissner A. 2010. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat. Biotechnol.* 28(5):521–26

Son EY, Ichida JK, Wainger BJ, Toma JS, Rafuse VF, et al. 2011. Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell* 9(3):205–18

Song K, Nam Y-J, Luo X, Qi X, Tan W, et al. 2012. Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature* 485(7400):599–604

Soufi A, Donahue G, Zaret KS. 2012. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* 151(5):994–1004

Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS. 2015. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* 161(3):555–68

Spemann H, Mangold H. 1923. Induction of embryonic primordia by implantation of organizers from a different species, transl. V Hamburger, 2001. *Int. J. Dev. Biol.* 45:13–38 (from German)

Stadtfeld M, Hochedlinger K. 2010. Induced pluripotency: history, mechanisms, and applications. *Genes Dev*. 5:582–96

Stefanakis N, Carrera I, Hobert O. 2015. Regulatory logic of pan-neuronal gene expression in *C. elegans*. *Neuron* 87(4):733–50

Sugimoto K, Gordon SP, Meyerowitz EM. 2011. Regeneration in plants and animals: dedifferentiation, transdifferentiation, or just differentiation? *Trends Cell Biol*. 21(4):212–18

Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T. 2001. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr. Biol.* 11(19):1553–58

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126(4):663–76

Tapscott SJ. 2005. The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. *Development* 132(12):2685–95

Taylor SM, Jones PA. 1979. Multiple new phenotypes induced in 10T12 and 3T3 cells treated with 5-azacytidine. *Cell* 17(4):771–79

Torper O, Ottosson DR, Pereira M, Lau S, Cardoso T, et al. 2015. In vivo reprogramming of striatal NG2 glia into functional neurons that integrate into local host circuitry. *Cell Rep*. 12(3):474–81

Torper O, Ulrich P, Wolf DA, Pereira M, Lau S, et al. 2013. Generation of induced neurons via direct conversion in vivo. *PNAS* 110(17):7038–43

Treutlein B, Lee Q, Camp GJ, Mall M, Koh W, et al. 2016. Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* 534(7607):391–95

Tsunemoto R, Lee S, Szűcs A, Chubukov P, Sokolova I, et al. 2018. Diverse reprogramming codes for neuronal identity. *Nature* 557(7705):375–80

Velasco S, Ibrahim MM, Kakumanu A, Garipler G, Aydin B, et al. 2017. A multi-step transcriptional and chromatin state cascade underlies motor neuron programming from embryonic stem cells. *Cell Stem Cell* 20(2):205–17.e8

Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Südhof TC, Wernig M. 2010. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463(7284):1035–41

Waddington CH. 1957. *The Strategy of the Genes*. London: George Allen Unwin

Wapinski OL, Vierbuchen T, Qu K, Lee QY, Chanda S, et al. 2013. Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* 155(3):621–35

Weintraub H, Tapscott SJ, Davis RL, Thayer MJ, Adam MA, et al. 1989. Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *PNAS* 86(14):5434–38

Wichterle H, Lieberam I, Porter JA, Jessell TM. 2002. Directed differentiation of embryonic stem cells into motor neurons. *Cell* 110(3):385–97

Yamanaka S. 2009. Elite and stochastic models for induced pluripotent stem cell generation. *Nature* 460:49–52

Yang N, Ng YH, Pang ZP, Südhof TC, Wernig M. 2011. Induced neuronal cells: how to make and define a neuron. *Cell Stem Cell* 9(6):517–25

Yoo AS, Sun AX, Li L, Shcheglovitov A, Portmann T, et al. 2011. MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature* 476:228–31

Zaret KS, Carroll JS. 2011. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev*. 25(21):2227–41

Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA. 2008. In vivo reprogramming of adult pancreatic exocrine cells to β-cells. *Nature* 455(7213):627–32