frontiers in Bioengineering and Biotechnology

MINI REVIEW
published: 18 October 2021
doi: 10.3389/fbioe.2021.748942

Single-Cell Genomics: Catalyst for Cell Fate Engineering

Boxun Li[1] *and Gary C. Hon*[1,2]*

[1]*Cecil H. and Ida Green Center for Reproductive Biology Sciences, University of Texas Southwestern Medical Center, Dallas, TX, United States,* [2]*Division of Basic Reproductive Biology Research, Department of Obstetrics and Gynecology, Department of Bioinformatics, University of Texas Southwestern Medical Center, Dallas, TX, United States*

As we near a complete catalog of mammalian cell types, the capability to engineer specific cell types on demand would transform biomedical research and regenerative medicine. However, the current pace of discovering new cell types far outstrips our ability to engineer them. One attractive strategy for cellular engineering is direct reprogramming, where induction of specific transcription factor (TF) cocktails orchestrates cell state transitions. Here, we review the foundational studies of TF-mediated reprogramming in the context of a general framework for cell fate engineering, which consists of: discovering new reprogramming cocktails, assessing engineered cells, and revealing molecular mechanisms. Traditional bulk reprogramming methods established a strong foundation for TF-mediated reprogramming, but were limited by their small scale and difficulty resolving cellular heterogeneity. Recently, single-cell technologies have overcome these challenges to rapidly accelerate progress in cell fate engineering. In the next decade, we anticipate that these tools will enable unprecedented control of cell state.

Keywords: reprogramming, single cell genomics, regenerative medicine, cell fate, transcription factor

OPEN ACCESS

Edited by:
Umberto Galderisi,
University of Campania Luigi Vanvitelli, Italy

Reviewed by:
Baris Tursun,
Max Delbrück Center for Molecular Medicine, Germany
Dario Siniscalco,
University of Campania Luigi Vanvitelli, Italy
Jose Polo,
Monash University, Australia

***Correspondence:**
Gary C. Hon
Gary.Hon@UTSouthwestern.edu

Specialty section:
This article was submitted to Preclinical Cell and Gene Therapy, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *16 August 2021*
Accepted: *05 October 2021*
Published: *18 October 2021*

Citation:
Li B and Hon GC (2021) Single-Cell Genomics: Catalyst for Cell Fate Engineering. Front. Bioeng. Biotechnol. 9:748942. doi: 10.3389/fbioe.2021.748942

INTRODUCTION

Progressive cell fate restriction is a central feature of organismal development famously illustrated by the "Waddington landscape" (Waddington, 1957). This model views cell fate establishment as irreversible. However, John Gurdon observed in *Xenopus* that nuclear transplantation of terminally differentiated cells into enucleated oocytes resulted in the development of normal frogs (Gurdon, 1962; Gurdon, 1967). This suggested that the nucleus does not permanently lose its potential to differentiate during development. In 1987, Davis and colleagues found that a gene specifically expressed in skeletal muscle, Myod1, converts mouse fibroblasts to skeletal muscle cells *in vitro* (Davis et al., 1987). In *Drosophila*, over-expression of the *eyeless* gene ectopically, eye structures are strikingly induced on the wings, the legs and the antennae (Halder et al., 1995). These studies clearly demonstrated the plasticity of terminally differentiated cells, and the possibility of engineering cell fate by gene over-expression. Two decades later, Takahashi and Yamanaka reprogrammed terminally differentiated cells to pluripotent stem cells with a cocktail of four transcription factors (TFs) (Takahashi and Yamanaka, 2006). This raised the important notion that cell fate engineering can be driven by a specific combination of TFs. Inspired by this breakthrough, many studies have extended this approach to reprogram pancreatic β-cells, cardiomyocytes, neurons, hepatocytes, and epicardial cells, among others ((Feng et al., 2008; Zhou et al., 2008; Ieda et al., 2010; Vierbuchen et al., 2010; Huang et al., 2011; Sekiya and Suzuki, 2011; Ladewig et al., 2012; Song et al., 2012; Nam et al., 2013; Niu et al., 2013; Batta et al., 2014; Chanda et al., 2014; Du et al., 2014; Riddell et al., 2014; Lemper et al., 2015; Duan et al., 2019) and reviewed in Xu et al. (2015), Wang et al. (2021).

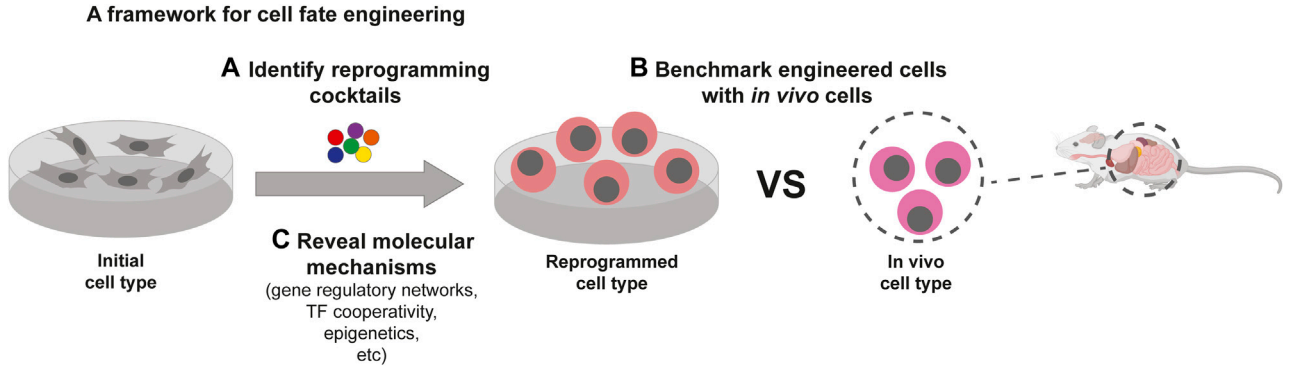

FIGURE 1 | Three key pillars of cell fate engineering. A generalizable framework for cell fate engineering will require: **(A)** the ability to discover new reprogramming cocktails at scale, **(B)** reliable ways to benchmark engineered cells with their endogenous counterparts, and **(C)** a deeper understanding of the underlying molecular mechanisms of cell fate conversion. Part of this figure was created with BioRender.com.

This capability to engineer cell fate holds great promise in regenerative medicine, disease modeling, and drug discovery (Grath and Dai, 2019; Zhang et al., 2020). Two strategies are most commonly used for engineering cell fate: A) direct engineering (used interchangeably with direct reprogramming), defined as conversion of cell fate without passing through an intermediate pluripotent state), and B) differentiation from a pluripotent state, e.g., induced pluripotent stem cells. Both are viable approaches with important differences and unique advantages. The comparison between the two strategies is beyond the scope of this review, and are discussed elsewhere (Margariti et al., 2014; Cieślar-Pobuda et al., 2017). In this review, we focus on direct engineering, though many principles discussed here can applied to the differentiation strategy.

Ideally, engineered cells need to faithfully recapitulate the target cell type at both molecular and functional levels. To extend cell fate engineering more broadly across cell types, tissues, and organisms, here we propose a methodological framework consisting of three pillars, based on current progress and future prospects of the field: 1) generalizable approaches to discovering new reprogramming cocktails at scale, 2) reliable ways to assess the engineered cells, benchmarked to their endogenous counterparts, and 3) comprehensive molecular mechanisms underlying cell fate engineering (**Figure 1**). Although stated separately, these areas of research are interrelated. For example, discovering new reprogramming cocktails usually involves some form of assessment of the engineered cells (Davis et al., 1987; Takahashi and Yamanaka, 2006; Zhou et al., 2008; Ieda et al., 2010; Vierbuchen et al., 2010; Song et al., 2012; Niu et al., 2013; Biddy et al., 2018; Duan et al., 2019); new molecular mechanisms often lead to improved reprogramming cocktails (Shu et al., 2013; Wapinski et al., 2013; Wang et al., 2015).

Despite recent progress, key challenges remain for each of the three pillars, which we will review here. First, we review the rationale and challenges of traditional methods that were developed for direct cell fate engineering before the advent of single-cell genomics. Second, we discuss promising single-cell genomic approaches that have emerged to address some of these challenges. Finally, we discuss the promise of applying this framework to *in vivo* reprogramming. Overall, we anticipate that single-cell approaches will play a key role in establishing a generalizable framework for cell fate engineering.

Traditional Methods: Rationale and Challenges

Pillar 1: Discovering Reprogramming Cocktails

Discovering cocktails of reprogramming factors is a two-step process. First, candidate genes must be selected. Several criteria are commonly used to identify candidate TFs that: 1) play a role in the natural development of target cell type, 2) manifest a relevant developmental phenotype when knocked out, and/or 3) are specifically expressed in the target cell type (Takahashi and Yamanaka, 2006; Ieda et al., 2010; Vierbuchen et al., 2010; Song et al., 2012). This curation step reduces the number of candidate genes to functionally test, which is critical to reduce the combinatorial space searched. However, these criteria also constrain these experiments to well-studied cell types.

Second, candidate genes are narrowed down to identify the smallest cocktail that efficiently reprograms cell fate. To efficiently achieve this goal, traditional approaches include the use of engineered reporters for successful cell fate conversion. Many studies rely on a single endogenous gene reporter that is engineered to be specifically expressed in the target cell type (Takahashi and Yamanaka, 2006; Ieda et al., 2010; Vierbuchen et al., 2010; Song et al., 2012). This is a useful way to simplify the readout and reduce the workload of screening through many cocktails. However, this convenience comes at the cost of two disadvantages: 1) a single reporter gene of successful reprogramming may not be known a priori or may not exist, and 2) the reporter may require the time-consuming task of genetically engineering cells or organisms. Another traditional approach used to increase efficiency, since the entire combinatorial space among all candidate genes is too prohibitive to search, is a 'minus one' experimental strategy that iteratively tests each factor's role in reprogramming by

removing it from the pool (Takahashi and Yamanaka, 2006; Ieda et al., 2010; Vierbuchen et al., 2010; Song et al., 2012). If the removal of a factor does not reduce or even increases the reprogramming efficiency, then it is deemed unnecessary and excluded from the pool. This is repeated multiple times until no factor can be subtracted without compromising reprogramming efficiency. This approach has two disadvantages. First, it is labor-intensive and tedious, making it hard to scale up to multiple target cell types. Second, this approach only searches a small proportion of the full combinatorial space among all candidate genes. This raises the possibility of missing alternative or more efficient cocktails. Indeed, Hand2 was shown to enhance the efficiency of the original GMT (Gata4, Mef2c, Tbx5) cocktail for cardiomyocyte reprogramming (Song et al., 2012). Intriguingly, in the screen conducted by Ieda and colleagues, removal of Hand2 increased reprogramming efficiency, and was thereby excluded from the cocktail (Ieda et al., 2010). This suggests that important TF interactions in reprogramming might not be readily revealed by the traditional screening approach.

Recent genomic strategies have addressed several shortcomings of traditional approaches. Two groups used CRISPR gene activation technology to screen through large numbers of putative TFs and other DNA-binding factors (2,428 and 1,496, respectively) for neuronal fate specification ability (Liu et al., 2018; Black et al., 2020). These scales are impressive, approximating the total number of all putative TFs in the human genome (Lambert et al., 2018). However, the studies have two limitations. First, they still rely on an endogenous reporter gene, the drawbacks of which have been discussed above. Second, since the screens were conducted in bulk, combinatorial perturbation information is lost. However, interaction within reprogramming cocktails is critical for cell fate engineering. For example, two studies screening hundreds of TF pairs for neuronal reprogramming performance revealed prevalent synergies between TFs that enhance reprogramming (Liu et al., 2018; Tsunemoto et al., 2018). While these studies highlight the importance of TF interactions, large scale screening of combinatorial TF cocktails remains challenging.

Recently, a suite of computational approaches has been developed to predict the reprogramming abilities of TFs and prioritize candidate TF cocktails to test experimentally (Cahan et al., 2014; Morris et al., 2014; D'Alessio et al., 2015; Rackham et al., 2016; Jung et al., 2021). Many of these methods rely on gene regulatory networks (GRNs) that link TFs to their target genes. These GRNs are often constructed from bulk gene expression datasets from diverse cell types and tissues, sometimes supplemented by bulk epigenetic data (Jung et al., 2021). These methods greatly reduce the combinatorial space of TF cocktails that need to be tested experimentally, addressing a major challenge posed above. Indeed, these methods have shown promising success in improving current reprogramming cocktails (Cahan et al., 2014; Morris et al., 2014), and predicting known (Rackham et al., 2016; Jung et al., 2021) and new (D'Alessio et al., 2015; Jung et al., 2021) reprogramming TFs. However, it remains to be seen if these methods are generalizable to a large number of target cell types. Moreover, one common drawback of these methods is that they often only use bulk, but not single-cell, expression and epigenetic datasets. As discussed more in detail below, bulk assays average across multiple cell types that often coexist in a given tissue, adding noise to cell type-specific GRN reconstruction that is key to the predictive power of these computational methods. As a result, future iterations of these methods should take advantage of the fast-expanding single-cell atlases to resolve the heterogeneity of bulk samples.

In summary, traditional strategies to discover reprogramming cocktails have advantages and disadvantages. Notably, the disadvantages stem from the lack of large-scale combinatorial screens, which limits both the scale and exhaustiveness of cocktail discovery. New technologies will be needed to address this challenge.

Pillar 2: Assessing Engineered Cells

Evaluating how well engineered cells recapitulate the molecular and functional features of endogenous cells is a critical task.

Molecular approaches are generalizable to different target cell types. Methods with simple readouts, such as real-time quantitative PCR and immunofluorescence, interrogate the changes of individual marker gene expression, usually targeted against well-established specific marker genes for the target cell type (Davis et al., 1987; Takahashi and Yamanaka, 2006; Zhou et al., 2008; Ieda et al., 2010; Vierbuchen et al., 2010; Song et al., 2012; Niu et al., 2013; Duan et al., 2019). While they are easy to implement, the expression of a handful of marker genes is hardly sufficient evidence of cell fate conversion. Thus, genome-scale readouts are often used to measure global changes in RNA or epigenetic (DNA methylation, chromatin accessibility, and histone marks) status in engineered cells (Takahashi and Yamanaka, 2006; Ieda et al., 2010; Wapinski et al., 2013; Cahan et al., 2014; Morris et al., 2014; Liu et al., 2016; Wapinski et al., 2017; Stone et al., 2019). These comprehensive molecular analyses have generated important insights. For example, by using bulk RNA microarray and ChIP-seq data to benchmark engineered cells across several target cell types, the CellNet studies illustrated that virtually all reprogramming paradigms fail to completely silence the gene expression programs of the starting cell (Cahan et al., 2014; Morris et al., 2014). While powerful, bulk genomic methods share a fundamental limitation: they take an average measurement of all cells in a population, thereby missing the heterogeneity and asynchrony of cell fate engineering (Treutlein et al., 2016; Liu et al., 2017; Biddy et al., 2018; Schiebinger et al., 2019; Zhou et al., 2019) (**Figure 2A**).

Functional assessment of engineered cells is vital for applications in regenerative medicine and disease modeling. This step entails testing the cellular functions most characteristic of and fundamental to the target cell type (Takahashi and Yamanaka, 2006; Zhou et al., 2008; Ieda et al., 2010; Vierbuchen et al., 2010; Song et al., 2012; Niu et al., 2013). For example, induced pluripotent stem cells are tested for their ability to differentiate into all three germ layers, as well as their contribution to mouse embryonic development (Takahashi and Yamanaka, 2006); induced pancreatic β-cells, their ability to

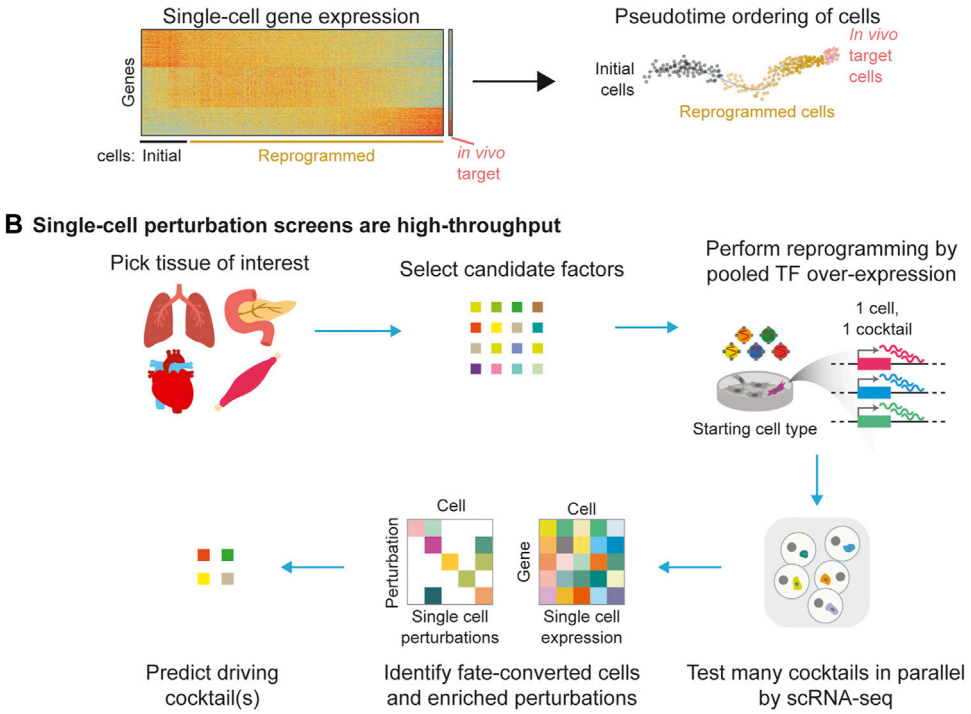

FIGURE 2 | Applications of single-cell genomics to cell fate engineering. **(A)** Single-cell analysis of reprogramming resolves cellular heterogeneity and reprogramming asynchrony. **(B)** Single-cell perturbation screens can scale the testing of reprogramming cocktails. Each sequenced cell tests a specific cocktail, and sequencing thousands of cells allows many cocktails to be simultaneously tested in one experiment.

secrete insulin and regulate blood glucose level (Zhou et al., 2008); induced cardiomyocytes, spontaneous beating and intracellular Ca^{2+} flux (Ieda et al., 2010). Furthermore, engineered cells need to survive engraftment and integrate with local cells *in vivo* (Ieda et al., 2010), if they are to be used for regenerative medicine. However, given that each target cell type has unique functions, functional assays are necessarily selected on an ad hoc basis, making it hard to generalize. As such, they are better reserved as the most stringent test, ideally performed after the cells pass the molecular tests mentioned above.

Pillar 3: Molecular Mechanisms

Here, we discuss two important molecular mechanisms of reprogramming: the combinatorial interactions of reprogramming TFs and the role of the epigenome.

Reprogramming cocktails usually contain multiple TFs, and many exhibit cooperative roles in reprogramming. First, some reprogramming TFs have pioneering activity (Cirillo et al., 2002; Soufi et al., 2012; Buganim et al., 2013; Wapinski et al., 2013), such as Gata4 in cardiomyocyte reprogramming (Ieda et al., 2010; Song et al., 2012), and Ascl1 in neuronal reprogramming (Vierbuchen et al., 2010). Pioneering is defined as the ability to bind regions of closed chromatin (Cirillo et al., 2002; Wapinski et al., 2013). It has been posited that pioneering TFs are essential to initiating cell fate engineering, but the maturation of product cells requires assistance from maturation factors, which is why the vast majority of reprogramming cocktails consists of at least one pioneering factor, plus other non-pioneering factors (Morris, 2016). Second, reprogramming TFs often interact to yield joint activities that are beyond individual TFs. For example, in neuronal reprogramming, Ascl1 binding recruits Brn2 to sites which are inaccessible to Brn2 alone (Wapinski et al., 2013). In cardiomyocyte reprogramming, Gata4, Mef2c, and Tbx5 cooperatively bind at cardiomyocyte-related genomic regions, and refine each other's binding affinity to certain regions when co-expressed (Stone et al., 2019). Moreover, Hand2 and Akt1 enhance the co-occupancy of GMT and GHMT at cardiomyocyte-related developmental enhancers, respectively (Hashimoto et al., 2019). Finally, the doses of TFs within a cocktail are important, shown by an elegant study using polycistronic constructs which found that GMT reprogramming is the most efficient when Mef2c is expressed more than the other two factors (Wang et al., 2015). This concept of a balance between reprogramming factors is further demonstrated by a seesaw model that argues that OCT4 and SOX2 counteract each other to achieve pluripotent reprogramming (Shu et al., 2013), among other reports that factor stoichiometry affects iPSC reprogramming efficiency and quality (Papapetrou et al., 2009; Carey et al., 2011; Tiemann et al., 2011; Yamaguchi et al., 2011). To sum up,

many important insights have been gained into the molecular mechanisms of reprogramming factors. With the discovery of more and more reprogramming cocktails, an important goal for future studies will be to determine the universality of these mechanisms to derive general rules of reprogramming.

Cell fate conversion requires both transcriptional (Takahashi and Yamanaka, 2006; Ieda et al., 2010; Cahan et al., 2014; Morris et al., 2014) and epigenetic reprogramming (Liu et al., 2016; Wapinski et al., 2017; Luo et al., 2019; Stone et al., 2019). Transcriptionally, reprogrammed cells activate target cell type-specific gene programs, and silence those from initial cell states (Takahashi and Yamanaka, 2006; Ieda et al., 2010). Yet, conversion often remains incomplete (Cahan et al., 2014; Morris et al., 2014). Epigenetically, chromatin accessibility, histone marks, and DNA methylation are globally reprogrammed (Liu et al., 2016; Wapinski et al., 2017; Luo et al., 2019; Stone et al., 2019). These are pivotal observations. However, one common drawback of the methods used in these studies is that they have most frequently been applied to bulk samples, which ignores the heterogeneity and asynchrony of reprogramming. For example, is incomplete transcriptional conversion (Cahan et al., 2014; Morris et al., 2014) a result of all cells being incompletely converted, or a mixture of fully and incompletely reprogrammed cells? Such questions can only be answered by single-cell analysis.

In summary, it is crucial to understand the functions and interaction of reprogramming factors, yet such knowledge is difficult to generate at large scale. Furthermore, our understanding of transcriptional and epigenetic changes that underlie cell fate engineering is limited by our ability to resolve heterogeneity in reprogramming. Single-cell approaches hold promise in addressing these gaps in knowledge.

Single-Cell Technologies: Promise and Limitations

Among the challenges discussed above, two themes are prominent and recurrent in traditional bulk reprogramming: resolution and scale (**Figure 2**). On one hand, bulk approaches cannot resolve the well-known heterogeneity in cell fate engineering. This hampers both the assessment of product cells and molecular mechanisms. On the other hand, bulk approaches also reduce the scale at which reprogramming cocktails can be tested. This limits the speed at which we can discover new cocktails and investigate molecular mechanisms. Excitingly, the advent of single-cell technologies have shown great promise in addressing these two challenges, albeit with their own limitations. Here, we discuss single-cell technologies with respect to the challenges of resolution and scale in cell fate engineering.

Single-Cell Omics Resolves Heterogeneity in Reprogramming

In the past decade, single-cell technologies have made great strides in lowering cost, increasing scale, and enabling new readouts (genomics, transcriptomics, and epigenomics) at unprecedented resolution (Telenius et al., 1992; Spits et al., 2006; Hashimshony et al., 2012; Ramsköld et al., 2012; Zong et al., 2012; Guo et al., 2013; Nagano et al., 2013; Picelli et al., 2013; Sasagawa et al., 2013; Smallwood et al., 2014; Buenrostro et al., 2015; Macosko et al., 2015; Rotem et al., 2015; Cao et al., 2017; Gierahn et al., 2017; Habib et al., 2017; Zheng et al., 2017; Han et al., 2018; Kaya-Okur et al., 2019; Ku et al., 2019), and reviewed in Kashima et al. (2020). These advances improve the assessment of engineered cells and aid the revelation of new molecular mechanisms by resolving heterogeneity (**Figure 2A**).

Single-cell RNA-sequencing (scRNA-seq) (Hashimshony et al., 2012; Ramsköld et al., 2012; Picelli et al., 2013; Sasagawa et al., 2013; Macosko et al., 2015; Cao et al., 2017; Gierahn et al., 2017; Habib et al., 2017; Zheng et al., 2017; Han et al., 2018) has been widely adopted in reprogramming studies and has led to several key observations. First, an early study of iPSC reprogramming analyzed 48 select genes of single cells during reprogramming (Buganim et al., 2012). The authors made an important observation that reprogramming is heterogeneous, with first a "stochastic" phase and then a more "deterministic" phase with Sox2 as the master regulator. This helps explain why only a small fraction of cells reach an iPSC fate. Second, a transcriptome-wide single-cell analysis of neuronal reprogramming induced by Ascl1, Brn2, and Myt1l strikingly revealed a reprogramming trajectory in which Ascl1 not only activates neuronal, but also myocytic, genes in fibroblasts, the latter of which are repressed by Brn2 and Myt1l (Treutlein et al., 2016). Therefore, cells that fail reprogramming might end up in an unproductive "dead-end" branch. Thirdly, a scRNA-seq analysis of human cardiomyocyte reprogramming induced by GMT revealed a decision point at which fibroblasts either progress further and become fully converted, or revert back to the fibroblast fate (Zhou et al., 2019). This raises the possibility that the previous observation of "incomplete" reprogramming (Cahan et al., 2014; Morris et al., 2014) is due to a mixture of successful and failed cells. Fourthly, scRNA-seq with dense time point sampling during iPSC reprogramming (Schiebinger et al., 2019) showed that "off target" cell fates are adopted, including stromal, trophoblast-like, and neuronal fates. Furthermore, the authors developed a computational method, Waddington-OT, that reconstructs the reprogramming trajectory, revealing a thin bottleneck for iPSC reprogramming. This conclusion was made possible because of the dense time points sampled, and demonstrates the power of using improved experimental designs to empower analytical methods. Moreover, they identified an environmental cue, GDF9, that is secreted by the stromal lineage to facilitate iPSC lineage reprogramming, echoing observations elsewhere that microenvironment plays an important role in reprogramming ((Mosteiro et al., 2016), also reviewed in Wang and Zhang (2018)). Finally, paralleling efforts on direct reprogramming, single-cell perturbation screens have also been applied to models of differentiation to resolve the functions of transcription factors during heterogeneous cell fate changes. Studies in definitive endoderm and teratoma differentiation delineate the TFs necessary for each cell state transition (Genga et al., 2019; McDonald et al., 2020).

Aside from the transcriptome, the measurement of other cellular features at single-cell resolution is still relatively immature. Nevertheless, some techniques have already been applied to cell fate engineering. First, single-cell

transcriptomics analysis combined with lineage tracing of fibroblast to endoderm progenitor reprogramming (Biddy et al., 2018) identified a successful path and a dead-end path. Interestingly, the cells commit to a path very early in reprogramming, when their global gene expression patterns have not yet diverged. Second, leveraging both scRNA-seq and single-cell ATAC-seq data, the same group developed CellOracle, a computational method to reconstruct GRNs which enabled them to identify additional factors at play in endoderm progenitor reprogramming (Kamimoto et al., 2020). Finally, with the development of new single-cell technologies that jointly measure multiple features in the same cell (Angermueller et al., 2016; Peterson et al., 2017; Stoeckius et al., 2017; Chen et al., 2019; Liu et al., 2019; Rooijers et al., 2019) and also reviewed in Kashima et al. (2020)), we anticipate that the transcriptomic and epigenomic changes in cell fate engineering will be increasingly dissected at the single-cell level, yielding broader and deeper insights.

Several limitations of single-cell technologies are worth noting. First, we and others observed that even at single-cell level, the most successfully reprogrammed cells are still not equivalent to their endogenous counterparts (Shin et al., 2012; Duan et al., 2019), at least at the transcriptional level. This could be due to incomplete maturation of product cells, the lack of the favorable microenvironment of the endogenous cells, or *in vitro* culture conditions. Second, transcriptional reprogramming does not necessarily dictate cellular functions. Therefore, functional validation is critical to complement single-cell omics assessment. Third, single-cell data is noisy (Adil et al., 2021), and the parameters of computational methods can influence data interpretation. Care should be taken to ensure that conclusions are robust across multiple parameters and algorithms.

Large-Scale Perturbation Screens Accelerate Cocktail Testing

Single-cell perturbation screens link each cell's transcriptome to its perturbation identity (Adamson et al., 2016; Dixit et al., 2016; Jaitin et al., 2016; Datlinger et al., 2017; Xie et al., 2017; Gasperini et al., 2019; Xie et al., 2019; Luginbühl et al., 2021). This methodology is powerful because it treats every sequenced cell as an independent experiment, thus measuring the transcriptional impact of genetic perturbations at scales previously unimaginable. For example, thousands of perturbations have been assayed in a recent study (Gasperini et al., 2019). In these screens, common types of perturbation are gene knockouts facilitated by genome editing (Dixit et al., 2016; Jaitin et al., 2016; Datlinger et al., 2017), gene knockdown by epigenome editing (i.e., CRISPR interference) (Adamson et al., 2016; Xie et al., 2017; Gasperini et al., 2019; Xie et al., 2019), and open reading frame over-expression of TFs (Parekh et al., 2018; Duan et al., 2019; Luginbühl et al., 2021). In the context of reprogramming, open reading frame over-expression is the most common mode of perturbation, used in all three single-cell perturbation studies (Parekh et al., 2018; Duan et al., 2019; Luginbühl et al., 2021), though gene activation and knockdown have also been used to achieve reprogramming (Liu et al., 2018; Qian et al., 2020; Zhou et al., 2020).

Before high-throughput scRNA-seq was widely available, an early study demonstrated the feasibility of screening for monocyte reprogramming factors at the single-cell level using multiplexed single-cell qPCR (Shin et al., 2012). Since only dozens of genes can be simultaneously measured in this way, the study required the assistance of the enrichment of established cell type-specific surface markers to report cell fate conversion. Applying modern single-cell perturbation screens to cell fate engineering, we and others have over-expressed pools of candidate factors and measured the joint readout of transcriptome and perturbation (Parekh et al., 2018; Duan et al., 2019; Luginbühl et al., 2021) of single cells (**Figure 2B**). This approach utilizes the transcriptomic readout to identify successfully reprogrammed cells, without relying on reporter genes. Then, the perturbation readout can be used to identify drivers of reprogramming. By undirected and directed combinatorial screening of 48 factors and 10 factors, respectively, we discovered Atf3, Gata6, and Hand2 to be a reprogramming cocktail for epicardial cells (Duan et al., 2019). Luginbühl and colleagues similarly identified previously unknown cocktails among 20 pro-neuronal TFs to engineer different neuronal subtypes (Luginbühl et al., 2021). Applying this concept to pluripotent stem cell differentiation models, Parekh and colleagues over-expressed 61 TFs and identified ETV2 as an inducer of endothelial fate (Parekh et al., 2018). These approaches are poised for expansion to more cell types. If successful, they will accelerate the discovery of new reprogramming cocktails.

Besides over-expression screens, repression screens of potential barriers of cell fate conversion are also of tremendous interest (Dhawan et al., 2011; Courtney et al., 2013; Tomaru et al., 2014; Zhou et al., 2016; Qian et al., 2020), also reviewed in Jenuwein and Allis (2001), Xu et al. (2015), Wang et al. (2021)). For example, the inhibition of epigenetic modifiers, such as histone deacetylases and polycomb complexes, and an RNA-binding protein, PTB, have been shown to facilitate or induce reprogramming (Xu et al., 2015; Zhou et al., 2016; Qian et al., 2020). In addition, the knock-down of four core regulators of dermal fibroblast cell fate enables adipogenesis under induction medium (Tomaru et al., 2014), suggesting that reprogramming can further benefit from disrupting combinations of factors in the starting cell type. In *c. elegans*, chromatin-regulating proteins including LIN-53, FACT, and MRG-1 have been shown to safeguard cell fate (Tursun et al., 2011; Kolundzic et al., 2018; Hajduskova et al., 2019). Knocking down these proteins facilitates cell fate conversion. Taken together, applying single-cell knock-down screens to existing master regulators or guardians of the starting cell fate could yield important insights into cell fate maintenance as well as powerful ways to facilitate reprogramming.

Single-cell combinatorial analysis of TFs can also yield new insights on molecular mechanisms. For example, this analysis can generate functional interaction data at large scale. Since single-cell combinatorial perturbation screens measure both the transcriptome and TF perturbation of a cell, single cells could be grouped computationally based on their perturbation. By

comparing the transcriptional effect of two single TFs, A and B, and that of both TFs together, AB, the interaction between A and B can be measured and modeled by established methods for analyzing genetic interactions, as described by Norman and colleagues (Norman et al., 2019). These mechanistic insights could be informative for the design of reprogramming cocktails.

In Vivo Cell Fate Engineering: Advantages and Challenges

In vivo cell fate engineering induces reprogramming factors *in vivo* (Zhou et al., 2008; Song et al., 2012; Niu et al., 2013), and presents unique advantages and challenges compared to *in vitro* reprogramming.

There are two important advantages. First, the *in vivo* microenvironment might be conducive to cell fate engineering, resulting in more efficient reprogramming and more mature product cells. For example, the three-factor cocktail that reprograms pancreatic exocrine cells to β-cells only works *in vivo*, but not *in vitro* (Zhou et al., 2008); cardiomyocyte reprogramming by GMT is more efficient *in vivo* than *in vitro* (Qian et al., 2012). Furthermore, neural injury and degeneration/aging have a positive impact on reprogramming (reviewed in Wang and Zhang (2018)). Second, *in vivo* cell fate engineering circumvents genetic mutations induced by *in vitro* culture, eliminating a major risk for regenerative medicine (Taguchi and Yamada, 2017).

However, there are also challenges associated with applying the methodological framework for cell fate engineering *in vivo*. First, *in vitro* findings of reprogramming cocktails do not always extrapolate to *in vivo* conditions ((Zhou et al., 2008), and reviewed in Tai et al. (2020)). Therefore, to achieve *in vivo* reprogramming, directly screening for cocktails *in vivo* is ideal. However, combinatorial gain-of-function screens *in vivo* are challenging to perform. As a result, *in vivo* gain-of-function TF screens have been limited to small numbers of candidate factors (Zhou et al., 2008; Niu et al., 2013). Therefore, there is an urgent need for scalable screening methods. Second, reprogrammed cells *in vivo* are embedded with endogenous cells, making it critical to distinguish engineered from non-engineered cells. Distinguishing these two kinds of cells is crucial to assess product cells and reveal molecular mechanisms. As such, extensive lineage tracing experiments are usually carried out in such studies (Zhou et al., 2008; Qian et al., 2012; Song et al., 2012; Grande et al., 2013; Niu et al., 2013; Guo et al., 2014).

CONCLUDING REMARKS

The field of cell fate engineering has made tremendous progress in the past 2 decades. In the 2000s, traditional bulk reprogramming approaches established a strong foundation for TF-mediated reprogramming, but were limited by cellular heterogeneity and small scale. In the 2010s, the development and application of single-cell technologies to cellular reprogramming has accelerated recent progress. In the coming decade, we anticipate that single-cell technologies will revolutionize the discovery of new reprogramming cocktails, the evaluation of engineered cells, and the molecular mechanisms of cell fate.

AUTHOR CONTRIBUTIONS

BL wrote and BL/GH jointly edited the manuscript.

ACKNOWLEDGMENTS

We acknowledge the following funding sources: CPRIT (RP190451), NIH (DP2GM128203, UM1HG011996), the Welch Foundation (I-1926-20170325), the Burroughs Wellcome Fund (1019804), and the Green Center for Reproductive Biology.

REFERENCES

Adamson, B., Norman, T. M., Jost, M., Cho, M. Y., Nuñez, J. K., Chen, Y., et al. (2016). A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell.* 167, 1867–1882. e21. doi:10.1016/j.cell.2016.11.048

Adil, A., Kumar, V., Jan, A. T., and Asger, M. (2021). Single-Cell Transcriptomics: Current Methods and Challenges in Data Acquisition and Analysis. *Front. Neurosci.* 15, 591122. doi:10.3389/fnins.2021.591122

Angermueller, C., Clark, S. J., Lee, H. J., Macaulay, I. C., Teng, M. J., Hu, T. X., et al. (2016). Parallel Single-Cell Sequencing Links Transcriptional and Epigenetic Heterogeneity. *Nat. Methods.* 13, 229–232. doi:10.1038/nmeth.3728

Batta, K., Florkowska, M., Kouskoff, V., and Lacaud, G. (2014). Direct Reprogramming of Murine Fibroblasts to Hematopoietic Progenitor Cells. *Cell Rep.* 9, 1871–1884. doi:10.1016/j.celrep.2014.11.002

Biddy, B. A., Kong, W., Kamimoto, K., Guo, C., Waye, S. E., Sun, T., et al. (2018). Single-Cell Mapping of Lineage and Identity in Direct Reprogramming. *Nature.* 564, 219–224. doi:10.1038/s41586-018-0744-4

Black, J. B., McCutcheon, S. R., Dube, S., Barrera, A., Klann, T. S., Rice, G. A., et al. (2020). Master Regulators and Cofactors of Human Neuronal Cell Fate Specification Identified by CRISPR Gene Activation Screens. *Cell Rep.* 33, 108460. doi:10.1016/j.celrep.2020.108460

Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., et al. (2015). Single-Cell Chromatin Accessibility Reveals Principles of Regulatory Variation. *Nature.* 523, 486–490. doi:10.1038/nature14590

Buganim, Y., Faddah, D. A., Cheng, A. W., Itskovich, E., Markoulaki, S., Ganz, K., et al. (2012). Single-Cell Expression Analyses During Cellular Reprogramming Reveal an Early Stochastic and a Late Hierarchic Phase. *Cell.* 150, 1209–1222. doi:10.1016/j.cell.2012.08.023

Buganim, Y., Faddah, D. A., and Jaenisch, R. (2013). Mechanisms and Models of Somatic Cell Reprogramming. *Nat. Rev. Genet.* 14, 427–439. doi:10.1038/nrg3473

Cahan, P., Li, H., Morris, S. A., Lummertz da Rocha, E., Daley, G. Q., and Collins, J. J. (2014). CellNet: Network Biology Applied to Stem Cell Engineering. *Cell.* 158, 903–915. doi:10.1016/j.cell.2014.07.020

Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., et al. (2017). Comprehensive Single-Cell Transcriptional Profiling of a Multicellular Organism. *Science* 357, 661–667. doi:10.1126/science.aam8940

Carey, B. W., Markoulaki, S., Hanna, J. H., Faddah, D. A., Buganim, Y., Kim, J., et al. (2011). Reprogramming Factor Stoichiometry Influences the Epigenetic State and Biological Properties of Induced Pluripotent Stem Cells. *Cell Stem Cell.* 9, 588–598. doi:10.1016/j.stem.2011.11.003

Chanda, S., Ang, C. E., Davila, J., Pak, C., Mall, M., Lee, Q. Y., et al. (2014). Generation of Induced Neuronal Cells by the Single Reprogramming Factor ASCL1. *Stem Cell Rep.* 3, 282–296. doi:10.1016/j.stemcr.2014.05.020

Chen, S., Lake, B. B., and Zhang, K. (2019). High-throughput Sequencing of the Transcriptome and Chromatin Accessibility in the Same Cell. *Nat. Biotechnol.* 37, 1452–1457. doi:10.1038/s41587-019-0290-0

Cieślar-Pobuda, A., Knoflach, V., Ringh, M. V., Stark, J., Likus, W., Siemianowicz, K., et al. (2017). Transdifferentiation and Reprogramming: Overview of the Processes, Their Similarities and Differences. *Biochim. Biophys. Acta Mol. Cell Res.* 1864, 1359–1369. doi:10.1016/j.bbamcr.2017.04.017

Cirillo, L. A., Lin, F. R., Cuesta, I., Friedman, D., Jarnik, M., and Zaret, K. S. (2002). Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4. *Mol. Cell.* 9, 279–289. doi:10.1016/s1097-2765(02)00459-8

Courtney, M., Gjernes, E., Druelle, N., Ravaud, C., Vieira, A., Ben-Othman, N., et al. (2013). The Inactivation of Arx in Pancreatic α-Cells Triggers Their Neogenesis and Conversion Into Functional β-like Cells. *Plos Genet.* 9, e1003934. doi:10.1371/journal.pgen.1003934

D'Alessio, A. C., Fan, Z. P., Wert, K. J., Baranov, P., Cohen, M. A., Saini, J. S., et al. (2015). A Systematic Approach to Identify Candidate Transcription Factors That Control Cell Identity. *Stem Cell Rep.* 5, 763–775. doi:10.1016/j.stemcr.2015.09.016

Datlinger, P., Rendeiro, A. F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., et al. (2017). Pooled CRISPR Screening With Single-Cell Transcriptome Readout. *Nat. Methods.* 14, 297–301. doi:10.1038/nmeth.4177

Davis, R. L., Weintraub, H., and Lassar, A. B. (1987). Expression of a Single Transfected cDNA Converts Fibroblasts to Myoblasts. *Cell.* 51, 987–1000. doi:10.1016/0092-8674(87)90585-x

Dhawan, S., Georgia, S., Tschen, S.-I., Fan, G., and Bhushan, A. (2011). Pancreatic β Cell Identity Is Maintained by DNA Methylation-Mediated Repression of Arx. *Dev. Cell.* 20, 419–429. doi:10.1016/j.devcel.2011.03.012

Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Arnon, L., et al. (2016). Perturb-Seq: Dissecting Molecular Circuits With Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell.* 167, 1853–1866. e17. doi:10.1016/j.cell.2016.11.038

Du, Y., Wang, J., Jia, J., Song, N., Xiang, C., Xu, J., et al. (2014). Human Hepatocytes With Drug Metabolic Function Induced From Fibroblasts by Lineage Reprogramming. *Cell Stem Cell.* 14, 394–403. doi:10.1016/j.stem.2014.01.008

Duan, J., Li, B., Bhakta, M., Xie, S., Zhou, P., Munshi, N. V., et al. (2019). Rational Reprogramming of Cellular States by Combinatorial Perturbation. *Cell Rep.* 27, 3486–3499. e6. doi:10.1016/j.celrep.2019.05.079

Feng, R., Desbordes, S. C., Xie, H., Tillo, E. S., Pixley, F., Stanley, E. R., et al. (2008). PU.1 and C/EBP/Convert Fibroblasts Into Macrophage-Like Cells. *Proc. Natl. Acad. Sci.* 105, 6057–6062. doi:10.1073/pnas.0711961105

Gasperini, M., Hill, A. J., McFaline-Figueroa, J. L., Martin, B., Kim, S., Zhang, M. D., et al. (2019). A Genome-Wide Framework for Mapping Gene Regulation via Cellular Genetic Screens. *Cell.* 176, 377–390. e19. doi:10.1016/j.cell.2018.11.029

Genga, R. M. J., Kernfeld, E. M., Parsi, K. M., Parsons, T. J., Ziller, M. J., and Maehr, R. (2019). Single-Cell RNA-Sequencing-Based CRISPRi Screening Resolves Molecular Drivers of Early Human Endoderm Development. *Cell Rep.* 27, 708–718. e10. doi:10.1016/j.celrep.2019.03.076

Gierahn, T. M., Wadsworth, M. H., 2nd, Hughes, T. K., Bryson, B. D., Butler, A., Satija, R., et al. (2017). Seq-Well: Portable, Low-Cost RNA Sequencing of Single Cells at High Throughput. *Nat. Methods.* 14, 395–398. doi:10.1038/nmeth.4179

Grande, A., Sumiyoshi, K., López-Juárez, A., Howard, J., Sakthivel, B., Aronow, B., et al. (2013). Environmental Impact on Direct Neuronal Reprogramming *In Vivo* in the Adult Brain. *Nat. Commun.* 4, 2373. doi:10.1038/ncomms3373

Grath, A., and Dai, G. (2019). Direct Cell Reprogramming for Tissue Engineering and Regenerative Medicine. *J. Biol. Eng.* 13, 14. doi:10.1186/s13036-019-0144-9

Guo, H., Zhu, P., Wu, X., Li, X., Wen, L., and Tang, F. (2013). Single-Cell Methylome Landscapes of Mouse Embryonic Stem Cells and Early Embryos Analyzed Using Reduced Representation Bisulfite Sequencing. *Genome Res.* 23, 2126–2135. doi:10.1101/gr.161679.113

Guo, Z., Zhang, L., Wu, Z., Chen, Y., Wang, F., and Chen, G. (2014). *In Vivo* direct Reprogramming of Reactive Glial Cells Into Functional Neurons after Brain Injury and in an Alzheimer's Disease Model. *Cell Stem Cell.* 14, 188–202. doi:10.1016/j.stem.2013.12.001

Gurdon, J. B. (1962). The Developmental Capacity of Nuclei Taken From Intestinal Epithelium Cells of Feeding Tadpoles. *J. Embryol. Exp. Morphol.* 10, 622–640. doi:10.1242/dev.10.4.622

Gurdon, J. B. (1967). On the Origin and Persistence of a Cytoplasmic State Inducing Nuclear DNA Synthesis in Frogs' Eggs. *Proc. Natl. Acad. Sci. U. S. A.* 58, 545–552. doi:10.1073/pnas.58.2.545

Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., et al. (2017). Massively Parallel Single-Nucleus RNA-Seq With DroNc-Seq. *Nat. Methods.* 14, 955–958. doi:10.1038/nmeth.4407

Hajduskova, M., Baytek, G., Kolundzic, E., Gosdschan, A., Kazmierczak, M., Ofenbauer, A., et al. (2019). MRG-1/MRG15 Is a Barrier for Germ Cell to Neuron Reprogramming in *Caenorhabditis elegans*. *Genetics.* 211, 121–139. doi:10.1534/genetics.118.301674

Halder, G., Callaerts, P., and Gehring, W. J. (1995). Induction of Ectopic Eyes by Targeted Expression of the Eyeless Gene in Drosophila. *Science.* 267, 1788–1792. doi:10.1126/science.7892602

Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., et al. (2018). Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell.* 173, 1307. doi:10.1016/j.cell.2018.05.012

Hashimoto, H., Wang, Z., Garry, G. A., Malladi, V. S., Botten, G. A., Ye, W., et al. (2019). Cardiac Reprogramming Factors Synergistically Activate Genome-Wide Cardiogenic Stage-Specific Enhancers. *Cell Stem Cell.* 25, 69. doi:10.1016/j.stem.2019.03.022

Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cell Rep.* 2, 666–673. doi:10.1016/j.celrep.2012.08.003

Huang, P., He, Z., Ji, S., Sun, H., Xiang, D., Liu, C., et al. (2011). Induction of Functional Hepatocyte-Like Cells From Mouse Fibroblasts by Defined Factors. *Nature.* 475, 386–389. doi:10.1038/nature10116

Ieda, M., Fu, J.-D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B. G., et al. (2010). Direct Reprogramming of Fibroblasts Into Functional Cardiomyocytes by Defined Factors. *Cell.* 142, 375–386. doi:10.1016/j.cell.2010.07.002

Jaitin, D. A., Weiner, A., Yofe, I., Lara-Astiaso, D., Keren-Shaul, H., David, E., et al. (2016). Dissecting Immune Circuits by Linking CRISPR-Pooled Screens With Single-Cell RNA-Seq. *Cell.* 167, 1883–1896. e15. doi:10.1016/j.cell.2016.11.039

Jenuwein, T., and Allis, C. D. (2001). Translating the Histone Code. *Science.* 293, 1074–1080. doi:10.1126/science.1063127

Jung, S., Appleton, E., Ali, M., Church, G. M., and Del Sol, A. (2021). A Computer-Guided Design Tool to Increase the Efficiency of Cellular Conversions. *Nat. Commun.* 12, 1659. doi:10.1038/s41467-021-21801-4

Kamimoto, K., Hoffmann, C. M., and Morris, S. A. (2020). CellOracle: Dissecting Cell Identity via Network Inference and In Silico Gene Perturbation. *bioRxiv*. doi:10.1101/2020.02.17.947416

Kashima, Y., Sakamoto, Y., Kaneko, K., Seki, M., Suzuki, Y., and Suzuki, A. (2020). Single-Cell Sequencing Techniques from Individual to Multiomics Analyses. *Exp. Mol. Med.* 52, 1419–1427. doi:10.1038/s12276-020-00499-2

Kaya-Okur, H. S., Wu, S. J., Codomo, C. A., Pledger, E. S., Bryson, T. D., Henikoff, J. G., et al. (2019). CUT&Tag for Efficient Epigenomic Profiling of Small Samples and Single Cells. *Nat. Commun.* 10, 1930. doi:10.1038/s41467-019-09982-5

Kolundzic, E., Ofenbauer, A., Bulut, S. I., Uyar, B., Baytek, G., Sommermeier, A., et al. (2018). FACT Sets a Barrier for Cell Fate Reprogramming in *Caenorhabditis elegans* and Human Cells. *Dev. Cell.* 46, 611–626. e12. doi:10.1016/j.devcel.2018.07.006

Ku, W. L., Nakamura, K., Gao, W., Cui, K., Hu, G., Tang, Q., et al. (2019). Single-Cell Chromatin Immunocleavage Sequencing (scChIC-Seq) to Profile Histone Modification. *Nat. Methods.* 16, 323–325. doi:10.1038/s41592-019-0361-7

Ladewig, J., Mertens, J., Kesavan, J., Doerr, J., Poppe, D., Glaue, F., et al. (2012). Small Molecules Enable Highly Efficient Neuronal Conversion of Human Fibroblasts. *Nat. Methods.* 9, 575–578. doi:10.1038/nmeth.1972

Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., et al. (2018). The Human Transcription Factors. *Cell.* 172, 650–665. doi:10.1016/j.cell.2018.01.029

Lemper, M., Leuckx, G., Heremans, Y., German, M. S., Heimberg, H., Bouwens, L., et al. (2015). Reprogramming of Human Pancreatic Exocrine Cells to β-Like Cells. *Cell Death Differ.* 22, 1117–1130. doi:10.1038/cdd.2014.193

Liu, L., Liu, C., Quintero, A., Wu, L., Yuan, Y., Wang, M., et al. (2019). Deconvolution of Single-Cell Multi-Omics Layers Reveals Regulatory Heterogeneity. *Nat. Commun.* 10, 470. doi:10.1038/s41467-018-08205-7

Liu, Y., Yu, C., Daley, T. P., Wang, F., Cao, W. S., Bhate, S., et al. (2018). CRISPR Activation Screens Systematically Identify Factors That Drive Neuronal Fate and Reprogramming. *Cell Stem Cell.* 23, 758–771. e8. doi:10.1016/j.stem.2018.09.003

Liu, Z., Chen, O., Zheng, M., Wang, L., Zhou, Y., Yin, C., et al. (2016). Re-Patterning of H3K27me3, H3K4me3 and DNA Methylation During Fibroblast Conversion Into Induced Cardiomyocytes. *Stem Cell Res.* 16, 507–518. doi:10.1016/j.scr.2016.02.037

Liu, Z., Wang, L., Welch, J. D., Ma, H., Zhou, Y., Vaseghi, H. R., et al. (2017). Single-cell Transcriptomics Reconstructs Fate Conversion From Fibroblast to Cardiomyocyte. *Nature.* 551, 100–104. doi:10.1038/nature24454

Luginbühl, J., Kouno, T., Nakano, R., Chater, T. E., Sivaraman, D. M., Kishima, M., et al. (2021). Decoding Neuronal Diversification by Multiplexed Single-Cell RNA-Seq. *Stem Cell Rep.* 16, 810–824. doi:10.1016/j.stemcr.2021.02.006

Luo, C., Lee, Q. Y., Wapinski, O., Castanon, R., Nery, J. R., Mall, M., et al. (2019). Global DNA Methylation Remodeling During Direct Reprogramming of Fibroblasts to Neurons. *Elife.* 8, e40197. doi:10.7554/eLife.40197

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., et al. (2015). Highly Parallel Genome-Wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell.* 161, 1202–1214. doi:10.1016/j.cell.2015.05.002

Margariti, A., Kelaini, S., and Cochrane, A. (2014). Direct Reprogramming of Adult Cells: Avoiding the Pluripotent State. *Stem Cells Cloning* 7, 19–29. doi:10.2147/sccaa.s38006

McDonald, D., Wu, Y., Dailamy, A., Tat, J., Parekh, U., Zhao, D., et al. (2020). Defining the Teratoma as a Model for Multi-Lineage Human Development. *Cell.* 183, 1–18. doi:10.1016/j.cell.2020.10.018

Morris, S. A., Cahan, P., Li, H., Zhao, A. M., San Roman, A. K., Shivdasani, R. A., et al. (2014). Dissecting Engineered Cell Types and Enhancing Cell Fate Conversion via CellNet. *Cell.* 158, 889–902. doi:10.1016/j.cell.2014.07.021

Morris, S. A. (2016). Direct Lineage Reprogramming via pioneer Factors; a Detour Through Developmental Gene Regulatory Networks. *Development.* 143, 2696–2705. doi:10.1242/dev.138263

Mosteiro, L., Pantoja, C., Alcazar, N., Marión, R. M., Chondronasiou, D., Rovira, M., et al. (2016). Tissue Damage and Senescence Provide Critical Signals for Cellular Reprogramming *In Vivo*. *Science.* 354, 467. doi:10.1126/science.aaf4445

Nagano, T., Lubling, Y., Stevens, T. J., Schoenfelder, S., Yaffe, E., Dean, W., et al. (2013). Single-Cell Hi-C Reveals Cell-To-Cell Variability in Chromosome Structure. *Nature.* 502, 59–64. doi:10.1038/nature12593

Nam, Y.-J., Song, K., Luo, X., Daniel, E., Lambeth, K., West, K., et al. (2013). Reprogramming of Human Fibroblasts Toward a Cardiac Fate. *Proc. Natl. Acad. Sci.* 110, 5588–5593. doi:10.1073/pnas.1301019110

Niu, W., Zang, T., Zou, Y., Fang, S., Smith, D. K., Bachoo, R., et al. (2013). *In Vivo* reprogramming of Astrocytes to Neuroblasts in the Adult Brain. *Nat. Cell Biol.* 15, 1164–1175. doi:10.1038/ncb2843

Norman, T. M., Horlbeck, M. A., Replogle, J. M., Ge, A. Y., Xu, A., Jost, M., et al. (2019). Exploring Genetic Interaction Manifolds Constructed from Rich Single-Cell Phenotypes. *Science.* 365, 786–793. doi:10.1126/science.aax4438

Papapetrou, E. P., Tomishima, M. J., Chambers, S. M., Mica, Y., Reed, E., Menon, J., et al. (2009). Stoichiometric and Temporal Requirements of Oct4, Sox2, Klf4, and C-Myc Expression for Efficient Human iPSC Induction and Differentiation. *Proc. Natl. Acad. Sci.* 106, 12759–12764. doi:10.1073/pnas.0904825106

Parekh, U., Wu, Y., Zhao, D., Worlikar, A., Shah, N., Zhang, K., et al. (2018). Mapping Cellular Reprogramming via Pooled Overexpression Screens With Paired Fitness and Single-Cell RNA-Sequencing Readout. *Cell Syst.* 7, 548–555. e8. doi:10.1016/j.cels.2018.10.008

Peterson, V. M., Zhang, K. X., Kumar, N., Wong, J., Li, L., Wilson, D. C., et al. (2017). Multiplexed Quantification of Proteins and Transcripts in Single Cells. *Nat. Biotechnol.* 35, 936–939. doi:10.1038/nbt.3973

Picelli, S., Björklund, Å. K., Faridani, O. R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for Sensitive Full-Length Transcriptome Profiling in Single Cells. *Nat. Methods.* 10, 1096–1098. doi:10.1038/nmeth.2639

Qian, H., Kang, X., Hu, J., Zhang, D., Liang, Z., Meng, F., et al. (2020). Reversing a Model of Parkinson's Disease With *In Situ* Converted Nigral Neurons. *Nature.* 582, 550–556. doi:10.1038/s41586-020-2388-4

Qian, L., Huang, Y., Spencer, C. I., Foley, A., Vedantham, V., Liu, L., et al. (2012). *In Vivo* Reprogramming of Murine Cardiac Fibroblasts Into Induced Cardiomyocytes. *Nature.* 485, 593–598. doi:10.1038/nature11044

Rackham, O. J. L., Firas, J., Fang, H., Oates, M. E., Holmes, M. L., Knaupp, A. S., et al. (2016). A Predictive Computational Framework for Direct Reprogramming Between Human Cell Types. *Nat. Genet.* 48, 331–335. doi:10.1038/ng.3487

Ramsköld, D., Luo, S., Wang, Y.-C., Li, R., Deng, Q., Faridani, O. R., et al. (2012). Full-Length mRNA-Seq from Single-Cell Levels of RNA and Individual Circulating Tumor Cells. *Nat. Biotechnol.* 30, 777–782. doi:10.1038/nbt.2282

Riddell, J., Gazit, R., Garrison, B. S., Guo, G., Saadatpour, A., Mandal, P. K., et al. (2014). Reprogramming Committed Murine Blood Cells to Induced Hematopoietic Stem Cells With Defined Factors. *Cell.* 157, 549–564. doi:10.1016/j.cell.2014.04.006

Rooijers, K., Markodimitraki, C. M., Rang, F. J., de Vries, S. S., Chialastri, A., de Luca, K. L., et al. (2019). Simultaneous Quantification of Protein–DNA Contacts and Transcriptomes in Single Cells. *Nat. Biotechnol.* 37, 766–772. doi:10.1038/s41587-019-0150-y

Rotem, A., Ram, O., Shoresh, N., Sperling, R. A., Goren, A., Weitz, D. A., et al. (2015). Single-Cell ChIP-Seq Reveals Cell Subpopulations Defined by Chromatin State. *Nat. Biotechnol.* 33, 1165–1172. doi:10.1038/nbt.3383

Sasagawa, Y., Nikaido, I., Hayashi, T., Danno, H., Uno, K. D., Imai, T., et al. (2013). Quartz-Seq: a Highly Reproducible and Sensitive Single-Cell RNA Sequencing Method, Reveals Non-genetic Gene-Expression Heterogeneity. *Genome Biol.* 14, R31. doi:10.1186/gb-2013-14-4-r31

Schiebinger, G., Shu, J., Tabaka, M., Cleary, B., Subramanian, V., Solomon, A., et al. (2019). Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming. *Cell.* 176, 928–943. e22. doi:10.1016/j.cell.2019.01.006

Sekiya, S., and Suzuki, A. (2011). Direct Conversion of Mouse Fibroblasts to Hepatocyte-Like Cells by Defined Factors. *Nature.* 475, 390–393. doi:10.1038/nature10263

Shin, J. W., Suzuki, T., Ninomiya, N., Kishima, M., Hasegawa, Y., Kubosaki, A., et al. (2012). Establishment of Single-Cell Screening System for the Rapid Identification of Transcriptional Modulators Involved in Direct Cell Reprogramming. *Nucleic Acids Res.* 40, e165. doi:10.1093/nar/gks732

Shu, J., Wu, C., Wu, Y., Li, Z., Shao, S., Zhao, W., et al. (2013). Induction of Pluripotency in Mouse Somatic Cells With Lineage Specifiers. *Cell.* 153, 963–975. doi:10.1016/j.cell.2013.05.001

Smallwood, S. A., Lee, H. J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J., et al. (2014). Single-Cell Genome-Wide Bisulfite Sequencing for Assessing Epigenetic Heterogeneity. *Nat. Methods.* 11, 817–820. doi:10.1038/nmeth.3035

Song, K., Nam, Y.-J., Luo, X., Qi, X., Tan, W., Huang, G. N., et al. (2012). Heart Repair by Reprogramming Non-Myocytes With Cardiac Transcription Factors. *Nature.* 485, 599–604. doi:10.1038/nature11139

Soufi, A., Donahue, G., and Zaret, K. S. (2012). Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement With the Genome. *Cell.* 151, 994–1004. doi:10.1016/j.cell.2012.09.045

Spits, C., Le Caignec, C., De Rycke, M., Van Haute, L., Van Steirteghem, A., Liebaers, I., et al. (2006). Whole-Genome Multiple Displacement Amplification from Single Cells. *Nat. Protoc.* 1, 1965–1970. doi:10.1038/nprot.2006.326

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P. K., Swerdlow, H., et al. (2017). Simultaneous Epitope and Transcriptome Measurement in Single Cells. *Nat. Methods.* 14, 865–868. doi:10.1038/nmeth.4380

Stone, N. R., Gifford, C. A., Thomas, R., Pratt, K. J. B., Samse-Knapp, K., Mohamed, T. M. A., et al. (2019). Context-specific Transcription Factor Functions Regulate Epigenomic and Transcriptional Dynamics During Cardiac Reprogramming. *Cell Stem Cell.* 25, 87–102. e9. doi:10.1016/j.stem.2019.06.012

Taguchi, J., and Yamada, Y. (2017). *In Vivo* reprogramming for Tissue Regeneration and Organismal Rejuvenation. *Curr. Opin. Genet. Dev.* 46, 132–140. doi:10.1016/j.gde.2017.07.008

Tai, W., Xu, X.-M., and Zhang, C.-L. (2020). Regeneration Through *In Vivo* Cell Fate Reprogramming for Neural Repair. *Front. Cell. Neurosci.* 14, 107. doi:10.3389/fncel.2020.00107

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells From Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell.* 126, 663–676. doi:10.1016/j.cell.2006.07.024

Telenius, H., Carter, N. P., Bebb, C. E., Nordenskjöld, M., Ponder, B. A., and Tunnacliffe, A. (1992). Degenerate Oligonucleotide-Primed PCR: General Amplification of Target DNA by a Single Degenerate Primer. *Genomics.* 13, 718–725. doi:10.1016/0888-7543(92)90147-k

Tiemann, U., Sgodda, M., Warlich, E., Ballmaier, M., Schöler, H. R., Schambach, A., et al. (2011). Optimal Reprogramming Factor Stoichiometry Increases Colony Numbers and Affects Molecular Characteristics of Murine Induced Pluripotent Stem Cells. *Cytometry A.* 79A, 426–435. doi:10.1002/cyto.a.21072

Tomaru, Y., Hasegawa, R., Suzuki, T., Sato, T., Kubosaki, A., Suzuki, M., et al. (2014). A Transient Disruption of Fibroblastic Transcriptional Regulatory Network Facilitates Trans-differentiation. *Nucleic Acids Res.* 42, 8905–8913. doi:10.1093/nar/gku567

Treutlein, B., Lee, Q. Y., Camp, J. G., Mall, M., Koh, W., Shariati, S. A. M., et al. (2016). Dissecting Direct Reprogramming From Fibroblast to Neuron Using Single-Cell RNA-Seq. *Nature.* 534, 391–395. doi:10.1038/nature18323

Tsunemoto, R., Lee, S., Szűcs, A., Chubukov, P., Sokolova, I., Blanchard, J. W., et al. (2018). Diverse Reprogramming Codes for Neuronal Identity. *Nature.* 557, 375–380. doi:10.1038/s41586-018-0103-5

Tursun, B., Patel, T., Kratsios, P., and Hobert, O. (2011). Direct Conversion of *C. elegans* Germ Cells Into Specific Neuron Types. *Science.* 331, 304–308. doi:10.1126/science.1199082

Vierbuchen, T., Ostermeier, A., Pang, Z. P., Kokubu, Y., Südhof, T. C., and Wernig, M. (2010). Direct Conversion of Fibroblasts to Functional Neurons by Defined Factors. *Nature.* 463, 1035–1041. doi:10.1038/nature08797

Waddington, C. H. (1957). "The Strategy of the Genes, a Discussion of Some Aspects of Theoretical Biology," in *With an Appendix [Some Physico-Chemical Aspects of Biological Organisation] by H. Kacser*. Editor C. H. Waddington (G. Allen and Unwin).

Wang, H., Yang, Y., Liu, J., and Qian, L. (2021). Direct Cell Reprogramming: Approaches, Mechanisms and Progress. *Nat. Rev. Mol. Cell Biol.* 22, 410. doi:10.1038/s41580-021-00335-z

Wang, L.-L., and Zhang, C.-L. (2018). Engineering New Neurons: *In Vivo* Reprogramming in Mammalian Brain and Spinal Cord. *Cell Tissue Res.* 371, 201–212. doi:10.1007/s00441-017-2729-2

Wang, L., Liu, Z., Yin, C., Asfour, H., Chen, O., Li, Y., et al. (2015). Stoichiometry of Gata4, Mef2c, and Tbx5 Influences the Efficiency and Quality of Induced Cardiac Myocyte Reprogramming. *Circ. Res.* 116, 237–244. doi:10.1161/circresaha.116.305547

Wapinski, O. L., Lee, Q. Y., Chen, A. C., Li, R., Corces, M. R., Ang, C. E., et al. (2017). Rapid Chromatin Switch in the Direct Reprogramming of Fibroblasts to Neurons. *Cell Rep.* 20, 3236–3247. doi:10.1016/j.celrep.2017.09.011

Wapinski, O. L., Vierbuchen, T., Qu, K., Lee, Q. Y., Chanda, S., Fuentes, D. R., et al. (2013). Hierarchical Mechanisms for Direct Reprogramming of Fibroblasts to Neurons. *Cell.* 155, 621–635. doi:10.1016/j.cell.2013.09.028

Xie, S., Armendariz, D., Zhou, P., Duan, J., and Hon, G. C. (2019). Global Analysis of Enhancer Targets Reveals Convergent Enhancer-Driven Regulatory Modules. *Cell Rep.* 29, 2570–2578. e5. doi:10.1016/j.celrep.2019.10.073

Xie, S., Duan, J., Li, B., Zhou, P., and Hon, G. C. (2017). Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. *Mol. Cell.* 66, 285–299. e5. doi:10.1016/j.molcel.2017.03.007

Xu, J., Du, Y., and Deng, H. (2015). Direct Lineage Reprogramming: Strategies, Mechanisms, and Applications. *Cell Stem Cell.* 16, 119–134. doi:10.1016/j.stem.2015.01.013

Yamaguchi, S., Hirano, K., Nagata, S., and Tada, T. (2011). Sox2 Expression Effects on Direct Reprogramming Efficiency as Determined by Alternative Somatic Cell Fate. *Stem Cell Res.* 6, 177–186. doi:10.1016/j.scr.2010.09.004

Zhang, Y., Xie, X., Hu, J., Afreen, K. S., Zhang, C.-L., Zhuge, Q., et al. (2020). Prospects of Directly Reprogrammed Adult Human Neurons for Neurodegenerative Disease Modeling and Drug Discovery: iN vs. iPSCs Models. *Front. Neurosci.* 14, 546484. doi:10.3389/fnins.2020.546484

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., et al. (2017). Massively Parallel Digital Transcriptional Profiling of Single Cells. *Nat. Commun.* 8, 14049. doi:10.1038/ncomms14049

Zhou, H., Su, J., Hu, X., Zhou, C., Li, H., Chen, Z., et al. (2020). Glia-to-Neuron Conversion by CRISPR-CasRx Alleviates Symptoms of Neurological Disease in Mice. *Cell.* 181, 590–603. e16. doi:10.1016/j.cell.2020.03.024

Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J., and Melton, D. A. (2008). *In Vivo* reprogramming of Adult Pancreatic Exocrine Cells to Beta-Cells. *Nature.* 455, 627–632. doi:10.1038/nature07314

Zhou, Y., Liu, Z., Welch, J. D., Gao, X., Wang, L., Garbutt, T., et al. (2019). Single-Cell Transcriptomic Analyses of Cell Fate Transitions During Human Cardiac Reprogramming. *Cell Stem Cell.* 25, 149. doi:10.1016/j.stem.2019.05.020

Zhou, Y., Wang, L., Vaseghi, H. R., Liu, Z., Lu, R., Alimohamadi, S., et al. (2016). Bmi1 Is a Key Epigenetic Barrier to Direct Cardiac Reprogramming. *Cell Stem Cell.* 18, 382–395. doi:10.1016/j.stem.2016.02.003

Zong, C., Lu, S., Chapman, A. R., and Xie, X. S. (2012). Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. *Science.* 338, 1622–1626. doi:10.1126/science.1229164

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Li and Hon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.