Mapping regulators of cell fate determination: Approaches and challenges

Aditya Kumar (iD), and Prashant Mali

View Online   Export Citation   CrossMark

ARTICLES YOU MAY BE INTERESTED IN

AIP Publishing

Mapping regulators of cell fate determination: Approaches and challenges

Aditya Kumar[a] (iD) and Prashant Mali[a]

AFFILIATIONS

Department of Bioengineering, University of California, San Diego, La Jolla, California 92093, USA

[a]Authors to whom correspondence should be addressed: adk001@eng.ucsd.edu and pmali@ucsd.edu

ABSTRACT

Given the limited regenerative capacities of most organs, strategies are needed to efficiently generate large numbers of parenchymal cells capable of integration into the diseased organ. Although it was initially thought that terminally differentiated cells lacked the ability to transdifferentiate, it has since been shown that cellular reprogramming of stromal cells to parenchymal cells through direct lineage conversion holds great potential for the replacement of post-mitotic parenchymal cells lost to disease. To this end, an assortment of genetic, chemical, and mechanical cues have been identified to reprogram cells to different lineages both *in vitro* and *in vivo*. However, some key challenges persist that limit broader applications of reprogramming technologies. These include: (1) low reprogramming efficiencies; (2) incomplete functional maturation of derived cells; and (3) difficulty in determining the typically multi-factor combinatorial recipes required for successful transdifferentiation. To improve efficiency by comprehensively identifying factors that regulate cell fate, large scale genetic and chemical screening methods have thus been utilized. Here, we provide an overview of the underlying concept of cell reprogramming as well as the rationale, considerations, and limitations of high throughput screening methods. We next follow with a summary of unique hits that have been identified by high throughput screens to induce reprogramming to various parenchymal lineages. Finally, we discuss future directions of applying this technology toward human disease biology via disease modeling, drug screening, and regenerative medicine.

I. INTRODUCTION

During development, cells become increasingly specialized to a terminally differentiated state. This highly regulated process is controlled in part by the expression of transcription factors (TFs) that form specific network modules to ensure stable gene expression and cell identity.[1,2] Maintaining cellular identity is critical for healthy organ function, leading some to theorize that terminally differentiated cells lack the ability to transdifferentiate to different lineages.[3] However, it has since been shown that the overexpression of only a few TFs can change cell identity.[4–6] For example, although there are estimated to be over 1500 different human TFs,[7] the overexpression of a single TF is capable of reprogramming fibroblasts into myoblasts.[8] Moreover, alternative combinations of TFs or other means of reprogramming, such as chemical and environmental cues, can lead to the same cell type, suggesting that the genetic regulatory networks characteristic of a specific cell type may be established by different reprogramming factors.[9–11] As such, enormous effort has been applied to identify these master regulators of cell fate. This concept has since attracted enormous interest in the regenerative medicine field as a means to regenerate diseased tissue by reprogramming stromal cells into post-mitotic parenchymal cells, yielding promising results across multiple lineages *in vitro* and even improving outcomes in *in vivo* models.[12–14]

However, low reprogramming efficiencies with known factors limit clinical translation.[15] Although only a few TFs are needed to induce reprogramming, the large set of potential combinations of TFs makes finding the optimal reprogramming cocktail for each cell type daunting.[16] Original screening methods were labor intensive and low throughput, but the recent application of high throughput genetic and chemical screening approaches has systematically identified new factors to improve efficiency.[17–20] In this review, we will discuss screening strategies utilized to reprogram cells toward parenchymal lineages. We will address considerations in interpreting results from screens while highlighting the strengths and weaknesses of different approaches. We will then highlight key factors found through screening approaches before concluding by discussing future directions of applying this technology toward human disease biology via disease modeling, drug screening, and regenerative medicine.

II. RATIONALE, CONSIDERATIONS, AND LIMITATIONS FOR USING HIGH THROUGHPUT SCREENS TO IDENTIFY REPROGRAMMING FACTORS

A. Rationale for utilizing high throughput screens to identify reprogramming factors

Although original strategies focused on the overexpression of a single TF to reprogram cells, it quickly became evident that specific combinations of multiple TFs are needed for efficient reprogramming of most cell types.[21,22] As a result, a trial and error approach was adopted where a pool of TFs thought to be important for lineage specification was delivered to cells and then a single TF was removed; if reprogramming efficiency went down, the factor was reinstated but if the efficiency went up or remained unchanged the factor was removed until a minimum combination was achieved.[23,24] A transformative success utilizing this approach was the reprogramming of fibroblasts into induced pluripotent stem cells (iPSCs) via overexpression of Oct4, Sox2, c-Myc, and Klf4.[25] 24 different genes associated with maintenance of embryonic stem cell identity were first overexpressed followed by removal of genes until the identification of the four factors needed to generate iPSCs.[25] This approach was rapidly adopted within the scientific community and applied to reprogram stromal cells into parenchymal cells of various lineages, including cardiomyocytes,[26] neurons,[27,28] hepatocytes,[29] and pancreatic islet cells.[30,31] Additionally, further work built upon this strategy by screening other genetic factors, such as microRNAs and chromatin modification,[32–35] as well as chemical reprogramming via the addition of cytokines and small molecule inhibitors.[36–40] While these pioneering studies definitively demonstrated the ability to directly reprogram stromal cells into parenchymal cells, this trial and error approach is limiting in the number of combinations that could be tested, resulting in low reprogramming efficiency.[41,42] As such, computational frameworks were developed to predict TF combinations that could directly reprogram one cell type into another.[43] While useful, these frameworks are based upon a similar concept of identifying TFs that are differentially expressed between cell types and serve more as a starting point in identifying potential combinations.

Given the potential number of TF combinations that are possible, high throughput screening approaches are critical for the unbiased identification of optimal TF combination. Recent technological advances have allowed researchers to move past the trial and error approach and instead combined high throughput genetic or chemical perturbations with phenotypic or transcriptomic readouts to identify factors governing cell fate (Table I).[16–19,35,44–61] These screens are either performed in arrayed format, where perturbations are maintained in separate culture conditions, or pooled format, where perturbations are assayed *en masse*. Although chemical screens must be run in an arrayed format, genetic screens are moving toward a pooled format. Both formats have their strengths and weakness. Pooled formats are easy to run and allow for comprehensive profiling of large libraries but require next generation sequencing (NGS) to map perturbations to phenotype and require single cell readout modalities. However, advances in NGS technology now make readout of pooled screen results quick and cheap, fueling the widespread use of pooled screens at a rapid pace. Arrayed formats require specialized automation to perform large screens, but it is very easy to correlate perturbations to phenotypes and to perform more rigorous phenotypic assays. Development of technology such as liquid handling robots and high content imaging allows for larger arrayed screens, but application is still limited due to

the cost of these setups (Table II). Regardless of how the screen is performed, they all rely on three key components: the screening format, the cell type/source, and the readout modality [Fig. 1(a)]. Consideration of the strengths and weakness of different methods for these components is critical to interpret results from high throughput screens and extract meaningful data.

B. Considerations in interpreting results from high throughput screens

1. Genetic and chemical screening formats

The two most common genetic screening formats performed are open reading frame (ORF) overexpression and clustered regularly interspaced short palindromic repeats (CRISPR) screens. Although ORF overexpression and CRISPR screens have similar goals, e.g., genetic perturbation, differences in how they modify gene expression lead to different strengths and weaknesses.[62] We will also briefly discuss RNA interference (RNAi) technology, which has played an important role in reprogramming screens[63] but has been largely phased out in favor of CRISPR screens.[64,65] Alternatively, chemical agents that target specific pathways or induce epigenetic remodeling can also serve as a means for cellular reprogramming. Several groups have performed high throughput chemical screens as an alternative or to augment genetic screens.[66] We will discuss the strengths and weakness of the various screening formats below as well as summarize these points in Table II.

ORF overexpression has classically been utilized to perform reprogramming studies.[6,67] This method involves cloning the entire coding sequence of a TF into a plasmid controlled by a promoter to drive exogenous gene expression. The plasmid is then packaged in a viral vector and delivered to the host cells so that the gene can be transcribed. The use of a promoter results in TFs being stably and dramatically overexpressed.[68] As would be expected, the degree of overexpression appears to be key to induce reprogramming; high overexpression of certain genes has a profound impact on the reprogramming efficiencies to iPSCs, where ten to twenty-fold increase in Oct4 and Klf4 gene expression led to higher quality fully reprogrammed iPSCs,[69] and cardiomyocytes, where a twofold additional increase in Mef2c overexpression resulted in a tenfold increase in reprogramming efficiency.[70] Furthermore, by defining the coding sequence, this approach allows for the expression of specific isoforms or mutant forms of a gene. This is important as specific isoforms may be needed to induce reprogramming.[71,72] The main disadvantage of ORF overexpression is the difficulty in scaling up libraries to perform a screen due to the large size of ORFs.[68] However, publicly available genome-scale ORFeome collections have been developed that contain validated sequences mapping to over 13 000 genes and can be rapidly swapped into a desired vector.[73]

The CRISPR/Cas9 complex was originally adapted to allow for the efficient deletion or insertion of DNA sequences into the genome. The complex is directed to its target site by a short single guide RNA (sgRNA) sequence that recognizes and binds to a specific DNA sequence in the genome. The Cas9 nuclease is then able to cut the DNA, preventing gene expression.[74,75] CRISPR knockout strategies have been used to reprogram cells via the inhibition of gene expression of TFs critical for lineage specification. For example, deletion of MyoD in C2C12 myoblasts led to transdifferentiation to adipogenic cells.[76]

TABLE I. Examples of reprogramming factors identified using high throughput screening.

Starting cell	Ending cell	Screen type	Readout	Unique hit	References
Mouse epiblast stem cell	iPSC	CRISPRa	Oct4-GFP	Sall1	19
Mouse embryonic fibroblast	iPSC	ORF overexpression	Nanog-GFP	Glis1	44
Mouse embryonic fibroblast	iPSC	ORF overexpression	Tra-1–60+	Hhex, Hlx	45
Mouse embryonic fibroblasts	iPSC	RNAi	Oct4-GFP	Trim28	46
Mouse embryonic fibroblast	iPSC	RNAi	Alkaline phosphatase+	Tox4	47
Mouse embryonic fibroblast	iPSC	Chemical	Oct4-GFP	AM580+EPZ004777+SGC0946 + 5-aza-2-deoxycitidine	48
Human dermal fibroblast	iPSC	Chemical	Alkaline phosphatase+	SAHA-PIP Ì	49
Mouse embryonic stem cell	Neuron	CRISPRa	Tubb3-CD8	Ezh2	18
Human fibroblasts	Neuron	ORF overexpression	scRNA-seq	Pax6+Neurog2+Dlx2+Zic1, Pax6+Neurog2 +Dlx1+Isl1	50
Mouse embryonic fibroblasts	Neuron	ORF overexpression	Tau-GFP	Brn3c+Ascl1	51
Mouse embryonic skin fibroblast	Neuron	Chemical	Tau-GFP	Forskolin+ISX9+ CHIR99021 +I-BET151	52
Human fetal lung fibroblasts	Neuron	Chemical	Map2+	Kenpaullone+Prostaglandin E2+Forskolin+BML210 + Aminoresveratolsulfat+PP2	53
Mouse cardiac fibroblast	Cardiac progenitor cell	CRISPR knockout	Nkx2-5-GFP	Dmap1	54
Mouse cardiac fibroblasts	Cardiomyocytes	ORF overexpression	αMHC-GFP	Znf281	55
Mouse cardiac fibroblasts	Cardiomyocytes	ORF overexpression	GCAMP+	Hand2+Nkx2.5+Gata4 +Mef2c+Tbx5	56
Mouse tail-tip fibroblasts	Cardiomyocytes	RNAi	αMHC-GFP	Bmi1	35
Mouse cardiac fibroblast	Cardiomyocytes	RNAi	αMHC-GFP	Zrsr2	57
Human cardiac fibroblasts	Cardiomyocytes	Chemical	Cardiac troponin T-GFP	SB431542+XAV939	58
Human iPSC	Endothelial cell	ORF overexpression	Fitness+ scRNA-seq	Etv2	17
Mouse embryonic fibroblast	Epicardial cell	ORF overexpression	scRNA-seq	Atf3+Gata6+Hand2	16
Mouse embryonic fibroblasts	Dendritic cell	ORF overexpression	Clec9a-tdTomato	Pu.1+Irf8+Batf3	59
Mouse embryonic stem cell	Primoridal germ cell	CRISPR knockout	Stella-GFP + Esg1-tdTomato	Nr5a2, Zfp296	60
Mouse embryonic stem cell	2C-like cell	CRISPRa	scRNA-seq	Dppa2, Smarca5	61

More recently, a catalytically dead form of Cas9 (dCas9) has been developed that can be fused to transcriptional agonists or antagonists to increase or decrease gene expression endogenously.[77–80] The main advantage of this approach is the ability to modulate gene expression without permanently modifying the genome. CRISPR activation (CRISPRa) induces gene expression through the fusion of dCas9 to transcriptional activators, such as VP64-p65-Rta, synergistic activation mediators, and SunTag, that have a DNA binding domain and a domain to activate transcription.[78–80] On the other hand, CRISPR interference (CRISPRi) suppresses gene expression through the fusion of dCas9 to transcriptional repressor peptides, such as KRAB, that sterically block transcriptional initiation or elongation.[81] CRISPR screens are popular due to the relative ease in developing large sgRNA libraries due to their small size, allowing for potentially genome wide perturbations.[82] The library size depends upon the target gene list, with multiple sgRNAs targeting each gene (typically 3–10). Correspondingly, genome-scale CRISPR screens have libraries ranging from 60 000 to 200 000 elements.[83] To modulate gene expression, sgRNAs are

TABLE II. Overview of various screening platforms.

	Equipment needed	Strengths	Weaknesses
CRISPR	1. Flow cytometer for reporter readout	1. Allows for endogenous gene overexpression	1. Off target risk, albeit small
	2. Next generation sequencers	2. Easy to develop large-scale library	2. Gene regulation not as strong as ORF systems
ORF	1. Flow cytometer for reporter readout	1. Strong overexpression	1. Difficult to make into large-scale library
	2. Next generation sequencers	2. Allows for expression of specific isoforms	2. Knockdown studies not readily feasible (require dominant negative versions of target protein coupled with strong overexpression)
RNAi	1. Flow cytometer for reporter readout	1. Allows for knockdown of critical genes	1. High off target risk impacts signal to noise in the screens
	2. Next generation sequencers	2. Easy to develop large-scale library	
Chemical	1. High content imaging system	1. Assay impact across a range of doses.	1. High throughput screening requires automation
	2. Liquid handling station	2. Easy to interpret results	2. Repeated dosing of small molecules

designed to target the promoter region or transcriptional start site of the gene.[84,85] Care must also be taken that sgRNAs are specific for their target sequence. Furthermore, some genes are controlled by multiple enhancers, which may result in inefficient gene activation or repression.[86,87]

In this regard, ORF overexpression approaches allow for strong overexpression of genes as well as the expression of specific isoforms or mutants, while CRISPRa allows for endogenous gene expression to more physiological levels and ease of use given that generating sgRNA libraries is much easier than ORF libraries. As such, these two approaches should be seen as complementary; indeed, a previous study comparing CRISPRa and ORF screens for resistance to MEK inhibitors noted that both screens shared a number of top hits while also identifying unique ones.[62] Given these factors, it is important to recognize the strengths and weaknesses when performing CRISPR or ORF overexpression screens.

For loss of function (LOF) studies, RNAi via the introduction of short-hairpin RNAs (shRNA) was commonly performed in the reprogramming field. Whereas ORF and CRISPR strategies target genomic DNA, RNAi promotes the degradation of complementary targeted mRNA. As such, RNAi results in a knockdown of gene expression but not a complete knockout.[65] This feature is useful when assessing developmentally critical genes, as often is performed in reprogramming screens, where a complete knockout could be lethal. In addition, the ease of creating a shRNA library made LOF screens an attractive alternative to ORF overexpression screens. Genome wide RNAi libraries can contain a similar number of elements to CRISPR libraries.[88] However, off-target effects are high as shRNAs can recognize and degrade mRNAs with imperfect complementary sequences,[89] potentially modulating expression of many transcripts and inducing phenotypes that can be dominant over intended behavior.[90] In a study that compared the on and off target effects of sgRNAs and shRNAs targeting the same genes, it was shown that 97.4% of sgRNAs had larger on target than off target effects, whereas only 41.8% of shRNAs had the same effect.[91] As a result, CRISPR-Cas based screening has

become an important approach in the field. The development of CRISPRi, which enables precision knockdown,[92] has also contributed to the growing toolset of complementary approaches for enabling functional genetic screening.

In addition to genetic screens, chemical screens that regulate cell signaling pathways or the activity of histone and DNA modifying enzymes have gained interest as an alternative or to supplement genetic screens in reprogramming studies.[93,94] The main advantage of a chemical approach is that it does not require the use of viral vectors, which reduces the risk of genomic instability or mutations. Furthermore, chemical compounds are more readily administered compared to the use of viral vectors and thus also more readily translatable.[93,94] Similar to genetic screens, chemical screens were first utilized to reprogram fibroblasts into iPSCs. Mouse embryonic fibroblasts were reprogrammed into iPSCs using a purely chemical cocktail composed of valproic acid, CHIR99021, RepSox, tranycypromine, forskolin, 3-deazaneplanocin A, and arotinoid acid.[95] This quickly led to the development of chemical cocktails capable of reprogramming stromal cells into various parenchymal cells. Chemical reprogramming shows clinical promise as well as it has recently been shown that *in vivo* reprogramming of fibroblasts to cardiomyocytes post-myocardial infarction is possible with a purely chemical cocktail, albeit involving the use of seven compounds.[96] Thus, similar to genetic reprogramming, the major issues associated with chemical reprogramming involve identifying compounds that are important for reprogramming while limiting the number of compounds necessary. However, high throughput chemical screening is much more labor intensive than genetic screening as chemical screens must be performed in an arrayed format as described earlier.[58,93,97] Furthermore, there are concerns of off-target effects unless better targeting chemical compounds can be produced.[98] Nonetheless, chemical screens offer an important method for cell reprogramming while further informing genetic screening targets by providing a better understanding of the signaling pathways involved in reprogramming.
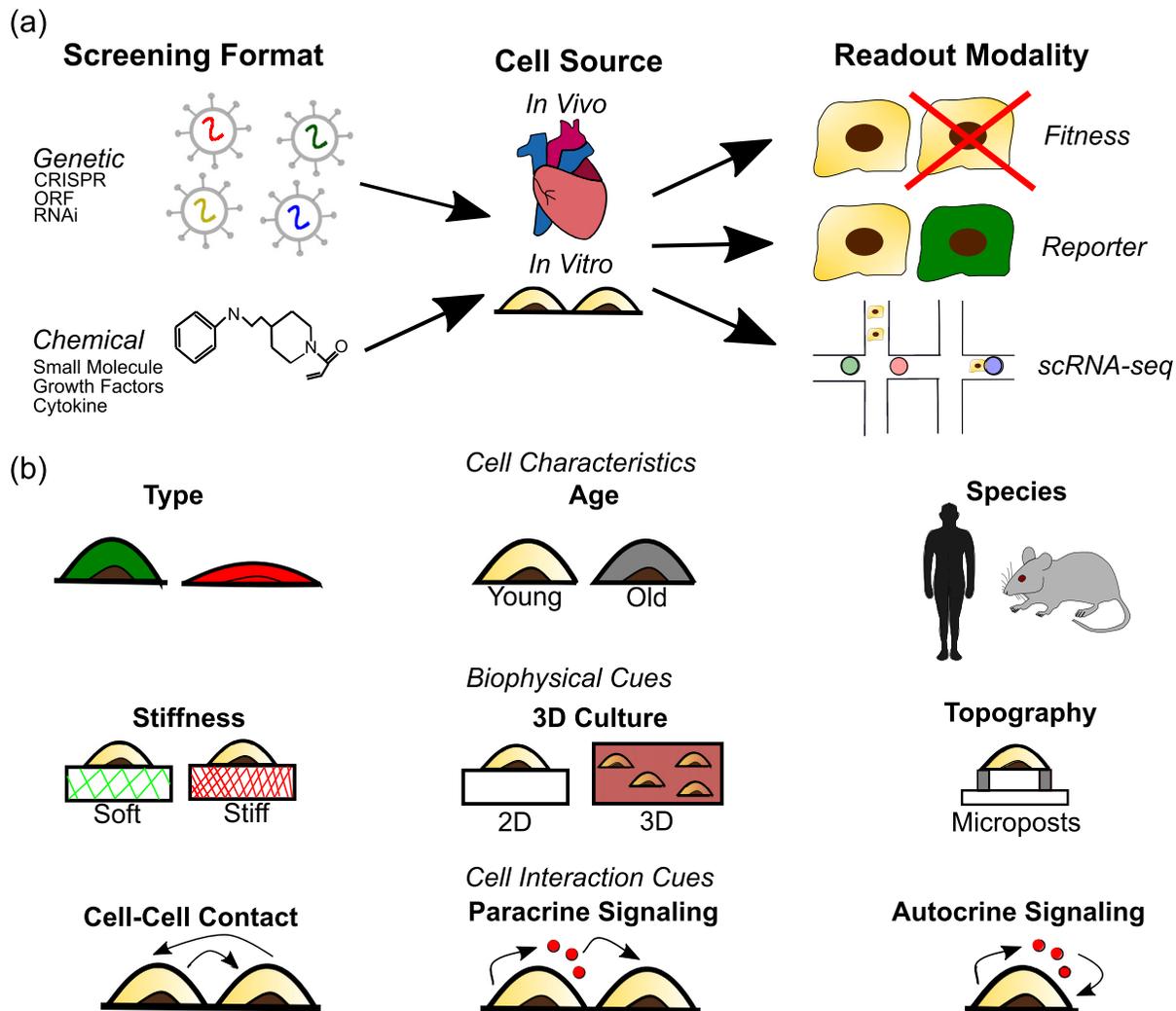
FIG. 1. Overview of the components and considerations involved in high throughput screening for the identification of reprogramming factors. (a) Listed are the common modalities utilized for reprogramming screens. (b) In addition to understanding the strengths and weakness of screen components, additional consideration of how cell characteristics, biophysical cues, and cell interaction cues can influence results is critical to properly interpret results from screens.

2. Readout modalities

Regardless of how the screen is performed, a phenotypic or transcriptomic readout is required to identify reprogrammed cells. The most common readout modalities utilized are fitness, reporter, or single cell RNA sequencing (scRNA-seq) readout.

Fitness readouts involve the enrichment or depletion of screening factors based on viability changes at the end of the screen.[86,99] To begin, cells are perturbed with the screen followed by additional culture to allow for the effects of perturbation to take place. Afterwards, genomic DNA is isolated from the remaining cell population and sequenced via NGS to map barcodes back to the starting library. Computational frameworks are used to determine differences in relative abundance of perturbations at the beginning and end of the screen to identify important factors. Top hits are then validated individually or in smaller screens and additional biological assays can then be performed to understand the mechanism by which enrichment or depletion occurs.[82,100] While these screens are simple to run, their application is limited as many cellular processes do not always involve changes in cellular viability, such as cell reprogramming.[101,102] However, fitness screens have been performed in the context of identifying factors responsible for the maintenance of stem cell or progenitor cell self-renewal[54,103] or the reprogramming of stem cells into parenchymal cells.[17] In this context, TFs that dropped out of the screen were investigated as potential reprogramming factors owing to the proliferation disadvantage associated with undergoing reprogramming without division.[17]

Reporter readouts, which involve the activation of a marker unique to a specific cell lineage, have classically been used to identify reprogrammed cell populations.[18,26,104] Generally, a fluorescent reporter gene is driven by a cell type specific promoter to allow for the

selection of reprogrammed cells via fluorescence activated cell sorting (FACS) after completion of the screen.[101] Common markers include green fluorescent protein under the control of alpha myosin heavy chain (αMHC) promoter for cardiomyocytes or beta-tubulin for neurons.[18,26] This readout is most commonly used for chemical screening due to its dependence on high content imaging to identify factors.[94,105] Although this readout modality has great applicability to reprogramming studies, there are some drawbacks. A unique marker only expressed by the reprogrammed cell must be identified, which may be difficult for every cell type. Furthermore, expression of the marker does not provide information on the global cellular transcriptome, thus limiting the ability to distinguish partially reprogrammed cells from fully reprogrammed ones.[106] For example, although a similar number of cells were αMHC positive for fibroblasts from different origins transduced with Gata4, Mef2c, and Tbx5, there were differences in the number of cells positive for cardiac troponin T between the groups.[26] Similar issues are associated with stem cell reprogramming, where partially reprogrammed clones may be positive for Nanog but not Oct4.[107] As such, any hit requires additional validation to separate false positives from confirmed factors.

scRNA-seq readouts involve the use of readable gRNAs or expressed barcodes that allow for mapping of phenotype changes to specific gene perturbations.[108,109] Whereas the previous screening approaches can assay for only simple phenotypes, such as viability or marker expression, scRNA-seq allows for screening of more complicated phenotypes and allows for a better understanding of the perturbation consequences. Several techniques, such as Perturb-Seq,[110] Crop-Seq,[111] CRISP-Seq,[112] and Mosaic-Seq,[113] have been developed that combine pooled barcoded CRISPR perturbations with scRNA-seq to assess the effects of single and multiple perturbations on the cell signaling response. In the reprogramming field, scRNA-seq has been utilized to understand transcriptomic differences that account for reprogramming heterogeneity, identifying barriers found in resistant cells that can be removed with additional factors.[114–117] Thus, scRNA-seq not only has utility as a readout but also as a tool to parse out the global effects of perturbation. Recently, this technology has been applied to high throughput reprogramming screens. Barcoded ORF overexpression libraries have been utilized to identify TFs responsible for reprogramming to endothelial[17] and epicardial-like[16] states based on their transcriptomic profiles. By combining TF overexpression with scRNA-seq, transcriptomic consequence of perturbations can be assessed, which allows for the identification of cell types as well as the separation of fully and partially reprogrammed cells. To successfully detect hits, the perturbation must have a strong effect or many cells carrying the specific perturbation must be sampled. Given the high sample preparation cost, profiling enough cells can be difficult. Furthermore, there are issues in separating runs due to sample-dependent batch effects.[114,118,119] However, there is significant interest in improving these weaknesses given the wealth of information gained through scRNA-seq. Improved computational techniques to reduce batch effects[120,121] and false positives[122,123] as well as techniques to allow for higher throughput through scRNA-seq multiplexing[124,125] have already been developed to improve data quality and reduce cost. Additional strategies, such as coupling scRNA-seq with other reporters to only sequence reprogrammed cells,[17] have been utilized to reduce cost and improve the confidence of scRNA-seq data.

It should be noted that the duration of screens is context dependent, in that the timing is dependent on the phenotype of interest and the time it takes for the genetic perturbation to be established and the resulting effect to manifest. In addition, duration is also readout dependent. Generally, *in vitro* assays involving fitness readouts take 2–4 weeks to allow for the enrichment of the population. Assays with reporter or scRNA-seq readouts are generally quicker than those with fitness readouts as they do not require population enrichment.[126]

3. Dependence on environmental conditions and starting cell populations

In addition to the components associated with the screen, it is important to understand how environmental conditions and starting cell populations can influence the results of a screen [Fig. 1(b)]. For example, it has been noted that *in vivo* reprogramming is both more efficient and induces cells with greater maturity than *in vitro* reprogramming with the same factors.[127] One explanation for this phenomenon is biophysical and biochemical differences between *in vivo* and *in vitro* extracellular environments.[128,129] Dynamic changes in the extracellular matrix occur during disease,[130,131] and numerous papers have demonstrated how mimicking these changing properties can induce changes in cell behavior.[132–134] These conditions are dramatically different from those in traditional reprogramming screens, in which cells are grown in plastic petri dishes. Topographical cues, such as microgrooves and micrograted substrates, have been shown to improve reprogramming of fibroblasts to iPSCs,[135] neurons,[136] and cardiomyocytes[137] via epigenetic modulation. However, responses to other cues appear to be cell type specific. Whereas matrix stiffness improves reprogramming efficiency to iPSCs and influences stem cell lineage specification, it had no effect on reprogramming to cardiomyocytes.[137–139] Instead, culture in 3D fibrin hydrogels improved cardiomyocyte reprogramming.[140] Biochemical cues via cell-cell contacts and paracrine or autocrine signaling have also been shown to influence reprogramming efficiency. For example, co-culture of stromal cells with the parenchymal cell of interest improves reprogramming efficiency, viability, and functionality.[141,142] Secreted factors can also influence reprogramming efficiency. It has been shown that blocking autocrine vascular endothelial growth factor signaling via knocking down vascular endothelial growth factor receptor 2 promotes stem cell self-renewal and somatic cell reprogramming.[143] Thus, it is important to note that possible transcriptomic changes driven by cell and environmental interactions rather than reprogramming factors can increase variability within the screen. This point is especially important in interpreting scRNA-seq data where the local environment of each cell is dependent on its interactions with neighboring cells.

In addition to environmental conditions, it is important to note that reprogramming cocktails are heavily dependent upon the starting cell population age, source, and species. Studies comparing reprogramming efficiencies of fibroblasts to iPSCs from young (2 months old) and old mice (>2 years old) found that the older cells exhibited a two-fold reduction in reprogramming efficiency.[144] Although this trend is not quite as clear in human cells, with high variance of reprogramming efficiency of fibroblasts to iPSCs from 12 individuals between the ages of 8–64 years but not correlated with age,[145] it has been noted that iPSCs generated from older patients have an increased number of mutations,[146] though this does not necessarily alter their

differentiation capability.[147] Cell source is an additional consideration that is shown to influence reprogramming responses to the same factors. For example, Gata4, Mef2c, and Tbx5 overexpression was most efficient for mouse cardiac fibroblast reprogramming to cardiomyocytes[26] but a separate study using mouse embryonic fibroblasts found that myocardin, Mef2c, and Tbx5 was the best combination and the addition of Gata4 was detrimental.[148] Reprogramming efficiencies using the same factors are also different for different cell sources.[26,149] These results are perhaps not surprising given that cells from different sources will have different epigenetic profiles that will alter their response to reprogramming factors. However, fibroblasts within the same organ are a heterogeneous population with different embryonic origins,[150,151] suggesting that even cells from the same source may respond differently to reprogramming factors. Finally, species-specific differences between mouse and human cells must be taken into consideration. Gata4, Mef2c, and Tbx5 overexpression is insufficient to reprogram human cardiac fibroblasts. Instead, the addition of Esrrg, Mesp1, myocardin, and Zfpm2 was needed for sarcomere formation, calcium transients, and action potentials.[58] Furthermore, whereas reprogramming of human cells to cardiomyocytes occurred anywhere from 4 to 11 weeks after transfection, mouse cell reprogrammed within two weeks.[26,152] This trend holds for other cell types as it is generally noted that human cells reprogram slower and with lower efficiency than mouse cells.[6] Given these results, reprogramming cocktails must be validated in human cells to be considered viable targets for regenerative medicine.

C. Limitations of high throughput screens

In Secs. II B 1 and II B 2, we have highlighted various screening methods and modalities and identified weakness associated with each one. However, there are additional limitations associated with all screens that must be considered.

First, it is difficult to recapitulate multilineage differentiations seen in normal development. Whereas stem cell differentiations often produce both the parenchymal and supporting cell populations,[153] genetic or chemical perturbations only produce a single or few cell types. As a result, additional cocktails must be discovered for each cell type. As mentioned above, co-culture can play an important role in reprogramming efficiency but is also important for cell maturation. For example, spontaneous neuronal activity and bursting behavior only occurs when reprogrammed neurons are co-cultured with astrocytes.[154]

Second, screens require a large number of cells to ensure good coverage of the library. To confidently separate true reprogramming hits from false positives, multiple cells must receive each combination of reprogramming factors. Given that combinations scale exponentially, screens must be designed in a manner that finds the balance between the number of factors and combinations in each cell.[126,155] This can be an issue when using primary cells with limited proliferative capacity or are difficult to transduce. Thus, a common approach is to use immortalized cells for the initial screen followed by primary cells for targeted screens.[126] There are obvious issues with this strategy given the high chance for false positive and negatives. In this context, ORF overexpression generally needs fewer cells per perturbation as it drives stronger phenotypic effects compared to CRISPR overexpression.[17] Regardless, it is important to consider how many factors and combinations are feasible given the starting cell population.

Third, reprogrammed cells must remain committed to the new lineage once the perturbation is removed to have any relevance in an *in vivo* setting where constant dosing is not possible. Thus, the endogenous cellular network must be rewired and stably reprogrammed. As a result, it is important to monitor cells over time to ensure no signs of de-differentiation. For example, previous studies have noted that reprogrammed cardiomyocytes can lose sarcomere formation and marker expression after a month in culture in serum-containing media but maintained identity when cultured in defined conditions.[156] Better understanding of the mechanism leading to loss of identity could potentially lead to the identification of factors that improve lineage commitment.

III. APPLICATIONS OF CELL REPROGRAMMING TO TISSUE ENGINEERING AND REGENERATIVE MEDICINE

Given the care needed to set up, execute, and interpret results from large-scale screens, it is important to understand the strengths and weakness of the different applications of cell reprogramming. There are numerous interesting biological questions associated with understanding the regulators of cell fate and also important applications of reprogramming technologies to tissue engineering and regenerative medicine. As such, researchers have applied reprogrammed cells for disease modeling or the replacement of diseased parenchymal tissue with varying success.

Similar to iPSC-differentiated progenies, directly reprogrammed cells have been utilized in tissue engineering applications. Tissue engineered blood vessels have been produced via reprogrammed endothelial and smooth muscle cells, resulting in increased survival of mice after transplantation of the graft containing both cell types compared to transplantation with only the decellularized vessel.[157] Additional applications of reprogrammed cells include drug screening, where induced hepatocytes exhibited similar drug metabolizing response and comparable toxicity prediction ability to primary human hepatocytes for 25 different compounds.[158] Finally, patient specific reprogrammed cells have also been created that faithfully recapitulate phenotypes associated with the disease. For example, reprogrammed neurons from a patient with Huntington's disease exhibited neuritic breakdown, abnormal neuritic branching, increased cell death, and aggregation of mutant huntingtin.[159] Similar recapitulation of deficits in neuron behavior associated with other neuronal diseases has also been demonstrated.[160] The biggest issue with utilizing directly reprogrammed cells for tissue engineering applications is the risk of exhausting the original cell population. Cells must be immortalized to ensure self-renewal, risking alteration in reprogrammed cell behavior. The generation of iPSCs, which are inherently self-renewing, does not run this risk. Furthermore, the application of reprogrammed cells to tissue engineering is heavily dependent on the reprogramming efficiency of the cells. Whereas iPSC differentiation efficiency to cardiomyocytes is routinely greater than 90% after two weeks,[161] reprogramming efficiencies are at best 50% after four weeks and involve a complicated assortment of genetic and chemical factors.[58] Additional research is needed to identify factors that improve efficiency to allow for applications of reprogrammed cells to tissue engineering.

Applications of reprogrammed cells for regenerative medicine are already ongoing and appear very promising. *In vivo* reprogramming has been reported to ameliorate the effects of diseases such as stroke,[14] diabetes,[13] and myocardial infarction.[12] There are many

advantages that make direct reprogramming an attractive approach for regenerating tissue. Reprogrammed cells integrate into the organ,[12–14] an important consideration given persistent issues with survival and engraftment with cell transplantation strategies.[162] In addition, direct reprogramming is cheaper, safer, and faster than stem cell differentiation as it skips the pluripotent and intermediate progenitor stages.[15] For example, *in vivo* reprogramming occurs within two months for most cell types, which is often the time needed to simply establish a stable iPSC line.[12–14] Finally, the overexpression of specific genes allows for control of final cell fate. For example, specific subtypes of certain cells, such as excitatory or inhibitory neurons, can be generated via the overexpression of different TFs.[163,164] However, there are limitations associated with direct reprogramming that must be overcome to translate this technology to the clinic. The use of viral vectors to deliver reprogramming factors carries the risk for insertional mutagenesis.[165] To reduce this risk, viral vectors that do not integrate into the host genome has been utilized. One of the more attractive options is delivery via adeno-associated virus (AAV) as it does not integrate into the genome, has a low immunogenicity profile and high transduction efficiency and specificity based on AAV capsid serotype, and is used already in the clinic.[166] Furthermore, additional research is attempted to further improve the therapeutic capacity of AAV vectors through the engineering of novel AAV capsid constructs. For example, recombinant AAV vectors have improved specificity and transduction efficiency, reducing the viral load needed for gene delivery.[167] Additionally, improvement of reprogramming efficiency will further reduce the time needed to produce enough reprogrammed cells to have a clinical benefit. Given reduced efficiency in reprogramming human cells *in vitro*,[6,26,152] it is likely that this trend will hold *in vivo*. Thus, improving delivery methods and identifying factors that improve efficiency are necessary for clinical translation of reprogramming technology.

IV. SUMMARY AND FUTURE OUTLOOK

Direct reprogramming has rapidly emerged as a powerful tool for generating post-mitotic parenchymal cells lost to disease. Although numerous genetic, chemical, and environmental cues have been identified to induce reprogramming, the ideal cocktails for each cell type remain elusive. To systematically identify new factors, high throughput genetic and chemical screening approaches have been utilized. These approaches have greatly aided our understanding of genetic interactions associated with cell fate and identified new factors responsible for reprogramming. Given the great promise already seen in direct reprogramming to regenerative medicine, improving reprogramming efficiency by the identification of better factors will aid in the translation of this technology to the clinic.

AUTHOR CONTRIBUTIONS

A.K. and P.M. conceived and wrote the manuscript.

ACKNOWLEDGMENTS

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

[1]T. Frum and A. Ralston, Trends Genet. **31**, 402 (2015).
[2]T. Graf and T. Enver, Nature **462**, 587 (2009).
[3]J. M. W. Slack, Nat. Rev. Genet. **3**, 889 (2002).
[4]S. A. Morris, Development **143**, 2696 (2016).
[5]M. Iwafuchi-Doi and K. S. Zaret, Development **143**, 1833 (2016).
[6]K. Takahashi and S. Yamanaka, Nat. Rev. Mol. Cell Biol. **17**, 183 (2016).
[7]J. M. Vaquerizas, S. K. Kummerfeld, S. A. Teichmann, and N. M. Luscombe, Nat. Rev. Genet. **10**, 252 (2009).
[8]R. L. Davis, H. Weintraub, and A. B. Lassar, Cell **51**, 987 (1987).
[9]S. Y. Wong, J. Soto, and S. Li, Curr. Opin. Chem. Eng. **15**, 95 (2017).
[10]A. E. Omole and A. O. J. Fakoya, PeerJ **2018**, 1.
[11]C. Yu, K. Liu, S. Tang, and S. Ding, Curr. Opin. Genet. Dev. **28**, 50 (2014).
[12]K. Miyamoto, M. Akiyama, F. Tamura, M. Isomi, H. Yamakawa, T. Sadahiro, N. Muraoka, H. Kojima, S. Haginiwa, S. Kurotsu, H. Tani, L. Wang, L. Qian, M. Inoue, Y. Ide, J. Kurokawa, T. Yamamoto, T. Seki, R. Aeba, H. Yamagishi, K. Fukuda, and M. Ieda, Cell Stem Cell **22**, 91 (2018).
[13]Y. Wang, C. Dorrell, W. E. Naugler, M. Heskett, P. Spellman, B. Li, F. Galivo, A. Haft, L. Wakefield, and M. Grompe, Mol. Ther. **26**, 1327 (2018).
[14]Y. C. Chen, N. X. Ma, Z. F. Pei, Z. Wu, F. H. Do-Monte, S. Keefe, E. Yellin, M. S. Chen, J. C. Yin, G. Lee, A. Minier-Toribio, Y. Hu, Y. T. Bai, K. Lee, G. J. Quirk, and G. Chen, Mol. Ther. **28**, 217 (2020).
[15]D. Srivastava and N. DeWitt, Cell **166**, 1386 (2016).
[16]J. Duan, B. Li, M. Bhakta, S. Xie, P. Zhou, N. V. Munshi, and G. C. Hon, Cell Rep. **27**, 3486 (2019).
[17]U. Parekh, Y. Wu, D. Zhao, A. Worlikar, N. Shah, K. Zhang, and P. Mali, Cell Syst. **7**, 548 (2018).
[18]Y. Liu, C. Yu, T. P. Daley, F. Wang, W. S. Cao, S. Bhate, X. Lin, C. Still, H. Liu, D. Zhao, H. Wang, X. S. Xie, S. Ding, W. H. Wong, M. Wernig, and L. S. Qi, Cell Stem Cell **23**, 758 (2018).
[19]J. Yang, S. S. Rajan, M. J. Friedrich, G. Lan, X. Zou, H. Ponstingl, D. A. Garyfallos, P. Liu, A. Bradley, and E. Metzakopian, Stem Cell Rep. **12**, 757 (2019).
[20]X. Xie, Y. Fu, and J. Liu, Curr. Opin. Genet. Dev. **46**, 104 (2017).
[21]C. F. Pereira, I. R. Lemischka, and K. Moore, Ann. N. Y. Acad. Sci. **1266**, 7 (2012).
[22]T. Ma, M. Xie, T. Laurent, and S. Ding, Circ. Res. **112**, 562 (2013).
[23]D. W. Allan and S. Thor, Wiley Interdiscip. Rev. Dev. Biol. **4**, 505 (2015).
[24]N. K. Wilson, S. D. Foster, X. Wang, K. Knezevic, J. Schütte, P. Kaimakis, P. M. Chilarska, S. Kinston, W. H. Ouwehand, E. Dzierzak, J. E. Pimanda, M. F. T. R. De Bruijn, and B. Göttgens, Cell Stem Cell **7**, 532 (2010).
[25]K. Takahashi and S. Yamanaka, Cell **126**, 663 (2006).
[26]M. Ieda, J. D. Fu, P. Delgado-Olguin, V. Vedantham, Y. Hayashi, B. G. Bruneau, and D. Srivastava, Cell **142**, 375 (2010).
[27]R. Ambasudhan, M. Talantova, R. Coleman, X. Yuan, S. Zhu, S. A. Lipton, and S. Ding, Cell Stem Cell **9**, 113 (2011).
[28]S. Chanda, C. E. Ang, J. Davila, C. Pak, M. Mall, Q. Y. Lee, H. Ahlenius, S. W. Jung, T. C. Südhof, and M. Wernig, Stem Cell Rep. **3**, 282 (2014).
[29]P. Huang, L. Zhang, Y. Gao, Z. He, D. Yao, Z. Wu, J. Cen, X. Chen, C. Liu, Y. Hu, D. Lai, Z. Hu, L. Chen, Y. Zhang, X. Cheng, X. Ma, G. Pan, X. Wang, and L. Hui, Cell Stem Cell **14**, 370 (2014).
[30]F. Thorel, V. Népote, I. Avril, K. Kohno, R. Desgraz, S. Chera, and P. L. Herrera, Nature **464**, 1149 (2010).
[31]E. Akinci, A. Banga, L. V. Greder, J. R. Dutton, and J. M. W. Slack, Biochem. J. **442**, 539 (2012).
[32]T. M. Jayawardena, B. Egemnazarov, E. A. Finch, L. Zhang, J. A. Payne, K. Pandya, Z. Zhang, P. Rosenberg, M. Mirotsou, and V. J. Dzau, Circ. Res. **110**, 1465 (2012).
[33]Y. Xue, K. Ouyang, J. Huang, Y. Zhou, H. Ouyang, H. Li, G. Wang, Q. Wu, C. Wei, Y. Bi, L. Jiang, Z. Cai, H. Sun, K. Zhang, Y. Zhang, J. Chen, and X. D. Fu, Cell **152**, 82 (2013).

[34]N. Singhal, J. Graumann, G. Wu, M. J. Araúzo-Bravo, D. W. Han, B. Greber, L. Gentile, M. Mann, and H. R. Schöler, Cell **141**, 943 (2010).

[35]Y. Zhou, L. Wang, H. R. Vaseghi, Z. Liu, R. Lu, S. Alimohamadi, C. Yin, J. D. Fu, G. G. Wang, J. Liu, and L. Qian, Cell Stem Cell **18**, 382 (2016).

[36]L. Baeyens, M. Lemper, G. Leuckx, S. De Groef, P. Bonfanti, G. Stangé, R. Shemer, C. Nord, D. W. Scheel, F. C. Pan, U. Ahlgren, G. Gu, D. A. Stoffers, Y. Dor, J. Ferrer, G. Gradwohl, C. V. E. Wright, M. Van De Casteele, M. S. German, L. Bouwens, and H. Heimberg, Nat. Biotechnol. **32**, 76 (2014).

[37]X. F. Zhao, J. Wan, C. Powell, R. Ramachandran, M. G. Myers, and D. Goldman, Cell Rep. **9**, 272 (2014).

[38]K. T. Lim, S. C. Lee, Y. Gao, K. P. Kim, G. Song, S. Y. An, K. Adachi, Y. J. Jang, J. Kim, K. J. Oh, T. H. Kwak, S. I. Hwang, J. S. You, K. Ko, S. H. Koo, A. D. Sharma, J. H. Kim, L. Hui, T. Cantz, H. R. Schöler, and D. W. Han, Cell Rep. **15**, 814 (2016).

[39]L. Zhang, J. C. Yin, H. Yeh, N. X. Ma, G. Lee, X. A. Chen, Y. Wang, L. Lin, L. Chen, P. Jin, G. Y. Wu, and G. Chen, Cell Stem Cell **17**, 735 (2015).

[40]N. Cao, Y. Huang, J. Zheng, C. I. Spencer, Y. Zhang, J.-D. Fu, B. Nie, M. Xie, M. Zhang, H. Wang, T. Ma, T. Xu, G. Shi, D. Srivastava, and S. Ding, Science **352**, 1216 (2016).

[41]S. Kelaini, A. Cochrane, and A. Margariti, Stem Cells Cloning Adv. Appl. **7**, 19 (2014).

[42]J. Hanna, K. Saha, B. Pando, J. Van Zon, C. J. Lengner, M. P. Creyghton, A. Van Oudenaarden, and R. Jaenisch, Nature **462**, 595 (2009).

[43]O. J. L. Rackham, J. Firas, H. Fang, M. E. Oates, M. L. Holmes, A. S. Knaupp, H. Suzuki, C. M. Nefzger, C. O. Daub, J. W. Shin, E. Petretto, A. R. R. Forrest, Y. Hayashizaki, J. M. Polo, and J. Gough, Nat. Genet. **48**, 331 (2016).

[44]M. Maekawa, K. Yamaguchi, T. Nakamura, R. Shibukawa, I. Kodanaka, T. Ichisaka, Y. Kawamura, H. Mochizuki, N. Goshima, and S. Yamanaka, Nature **474**, 225 (2011).

[45]T. Yamakawa, Y. Sato, Y. Matsumura, Y. Kobayashi, Y. Kawamura, N. Goshima, S. Yamanaka, and K. Okita, Stem Cells **34**, 2661 (2016).

[46]D. Miles, N. Alexandra de Vries, S. Gisler, C. Lieftink, W. Akhtar, E. Gogola, I. Pawlitzky, D. Hulsman, E. Tanger, M. Koppens, R. Beijersbergen, and M. Van Lohuizen, Stem Cells **35**, 147 (2017).

[47]L. Vanheer, J. Song, N. De Geest, A. Janiszewski, I. Talon, C. Provenzano, T. Oh, J. Chappell, and V. Pasque, J. Cell Sci. **132**, jcs232223 (2019).

[48]Y. Zhao, T. Zhao, J. Guan, X. Zhang, Y. Fu, J. Ye, J. Zhu, G. Meng, J. Ge, S. Yang, L. Cheng, Y. Du, C. Zhao, T. Wang, L. Su, W. Yang, and H. Deng, Cell **163**, 1678 (2015).

[49]G. N. Pandian, S. Sato, C. Anandhakumar, J. Taniguchi, K. Takashima, J. Syed, L. Han, A. Saha, T. Bando, H. Nagase, and H. Sugiyama, ACS Chem. Biol. **9**, 2729 (2014).

[50]J. Luginbuhl, T. Kouno, R. Nakano, T. Chater, D. Sivaraman, M. Kishima, F. Roudnicky, P. Carninci, C. Plessy, and J. Shin, biorxiv:600239 (2019).

[51]R. Tsunemoto, S. Lee, A. Szucs, P. Chubukov, I. Sokolova, J. W. Blanchard, K. T. Eade, J. Bruggemann, C. Wu, A. Torkamani, P. P. Sanna, and K. K. Baldwin, Nature **557**, 375 (2018).

[52]X. Li, X. Zuo, J. Jing, Y. Ma, J. Wang, D. Liu, J. Zhu, X. Du, L. Xiong, Y. Du, J. Xu, X. Xiao, J. Wang, Z. Chai, Y. Zhao, and H. Deng, Cell Stem Cell **17**, 195 (2015).

[53]U. Pfisterer, F. Ek, S. Lang, S. Soneji, R. Olsson, and M. Parmar, Sci. Rep. **6**, 38290 (2016).

[54]J. Yu, G. Palano, C. Lim, A. Moggio, L. Drowley, A. Plowright, M. Bohlooly, B. Rosen, E. Hansson, Q.-D. Wang, and K. Yusa, Stem Cells **37**, 958 (2019).

[55]H. Zhou, M. G. Morales, H. Hashimoto, M. E. Dickson, K. Song, W. Ye, M. S. Kim, H. Niederstrasser, Z. Wang, B. Chen, B. A. Posner, R. Bassel-Duby, and E. N. Olson, Genes Dev. **31**, 1770 (2017).

[56]R. C. Addis, J. L. Ifkovits, F. Pinto, L. D. Kellam, P. Esteso, S. Rentschler, N. Christoforou, J. A. Epstein, and J. D. Gearhart, J. Mol. Cell. Cardiol. **60**, 97 (2013).

[57]Y. Zhou, S. Alimohamadi, L. Wang, Z. Liu, J. B. Wall, C. Yin, J. Liu, and L. Qian, Stem Cells Int. **2018**, 6.

[58]T. M. A. Mohamed, N. R. Stone, E. C. Berry, E. Radzinsky, Y. Huang, K. Pratt, Y. S. Ang, P. Yu, H. Wang, S. Tang, S. Magnitsky, S. Ding, K. N. Ivey, and D. Srivastava, Circulation **135**, 978 (2017).

[59]F. F. Rosa, C. F. Pires, I. Kurochkin, A. G. Ferreira, A. M. Gomes, L. G. Palma, K. Shaiv, L. Solanas, C. Azenha, D. Papatsenko, O. Schulz, C. R. E. Sousa, and C. F. Pereira, Sci. Immunol. **3**, eaau4292 (2018).

[60]J. A. Hackett, Y. Huang, U. Günesdogan, K. A. Gretarsson, T. Kobayashi, and M. A. Surani, Nat. Commun. **9**, 5328 (2018).

[61]C. Alda-Catalinas, D. Bredikhin, I. Hernando-Herraez, O. Kubinyecz, F. Santos, M. Eckersley-Maslin, O. Stegle, and W. Reik, biorxiv:741371 (2019).

[62]K. R. Sanson, R. E. Hanna, M. Hegde, K. F. Donovan, C. Strand, M. E. Sullender, E. W. Vaimberg, A. Goodale, D. E. Root, F. Piccioni, and J. G. Doench, Nat. Commun. **9**, 5416 (2018).

[63]R. A. Panepucci and I. M. De Souza Lima, Stem Cell Res. Ther. **10**, 1 (2019).

[64]U. Unniyampurath, R. Pilankatta, and M. N. Krishnan, Int. J. Mol. Sci. **17**, 291 (2016).

[65]M. Boettcher and M. McManus, Mol. Cell **58**, 575 (2015).

[66]X. Ma, L. Kong, and S. Zhu, Protein Cell **8**, 328 (2017).

[67]F. González, S. Boué, and J. C. I. Belmonte, Nat. Rev. Genet. **12**, 231 (2011).

[68]D. Škalamera, M. V. Ranall, B. M. Wilson, P. Leo, A. S. Purdon, C. Hyde, E. Nourbakhsh, S. M. Grimmond, S. C. Barry, B. Gabrielli, and T. J. Gonda, PLoS One **6**, e20057 (2011).

[69]B. W. Carey, M. Markoulaki, J. H. Hanna, D. A. Faddah, Y. Buganim, J. Kim, K. Ganz, E. J. Steine, J. P. Cassady, M. P. Creyghton, G. G. Welstead, Q. Gao, and R. Jaenisch, Cell Stem Cell **9**, 588 (2011).

[70]L. Wang, Z. Liu, C. Yin, H. Asfour, O. Chen, Y. Li, N. Bursac, J. Liu, and L. Qian, Circ. Res. **116**, 237 (2015).

[71]Y. Tang, Z. Jiang, Y. Luo, X. Zhao, L. Wang, C. Norris, and X. C. Tian, J. Cell Sci. **127**, 3998 (2014).

[72]M. Gabut, P. Samavarchi-Tehrani, X. Wang, V. Slobodeniuc, D. O'Hanlon, H. K. Sung, M. Alvarez, S. Talukder, Q. Pan, E. O. Mazzoni, S. Nedelec, H. Wichterle, K. Woltjen, T. R. Hughes, P. W. Zandstra, A. Nagy, J. L. Wrana, and B. J. Blencowe, Cell **147**, 132 (2011).

[73]X. Yang, J. S. Boehm, X. Yang, K. Salehi-Ashtiani, T. Hao, Y. Shen, R. Lubonja, S. R. Thomas, O. Alkan, T. Bhimdi, T. M. Green, C. M. Johannessen, S. J. Silver, C. Nguyen, R. R. Murray, H. Hieronymus, D. Balcha, C. Fan, C. Lin, L. Ghamsari, M. Vidal, W. C. Hahn, D. E. Hill, and D. E. Root, Nat. Methods **8**, 659 (2011).

[74]L. Cong, F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. Hsu, X. Wu, W. Jiang, L. Marraffini, and F. Zhang, Science **339**, 819 (2013).

[75]P. Mali, L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, and G. M. Church, Science **339**, 823 (2013).

[76]C. Wang, W. Liu, Y. Nie, M. Qaher, H. E. Horton, F. Yue, A. Asakura, and S. Kuang, EBioMedicine **16**, 212 (2017).

[77]L. A. Gilbert, M. H. Larson, L. Morsut, Z. Liu, G. A. Brar, S. E. Torres, N. Stern-Ginossar, O. Brandman, E. H. Whitehead, J. A. Doudna, W. A. Lim, J. S. Weissman, and L. S. Qi, Cell **154**, 442 (2013).

[78]S. Chakraborty, H. Ji, A. M. Kabadi, C. A. Gersbach, N. Christoforou, and K. W. Leong, Stem Cell Rep. **3**, 940 (2014).

[79]D. Balboa, J. Weltner, S. Eurola, R. Trokovic, K. Wartiovaara, and T. Otonkoski, Stem Cell Rep. **5**, 448 (2015).

[80]J. Weltner, D. Balboa, S. Katayama, M. Bespalov, K. Krjutškov, E. M. Jouhilahti, R. Trokovic, J. Kere, and T. Otonkoski, Nat. Commun. **9**, 2643 (2018).

[81]R. M. J. Genga, E. M. Kernfeld, K. M. Parsi, T. J. Parsons, M. J. Ziller, and R. Maehr, Cell Rep. **27**, 708 (2019).

[82]E. H. Yau and T. M. Rana, Methods Mol. Biol. **1712**, 203 (2018).

[83]J. R. McDade, N. C. Waxmonsky, L. E. Swanson, and M. Fan, Curr. Protoc. Mol. Biol. **2016**, 31.5.1.

[84]M. Kampmann, ACS Chem. Biol. **13**, 406 (2018).

[85]B. P. Kleinstiver, M. S. Prew, S. Q. Tsai, V. V. Topkar, N. T. Nguyen, Z. Zheng, A. P. W. Gonzales, Z. Li, R. T. Peterson, J. R. J. Yeh, M. J. Aryee, and J. K. Joung, Nature **523**, 481 (2015).

[86]S. E. Mohr, Y. Hu, B. Ewen-Campen, B. E. Housden, R. Viswanatha, and N. Perrimon, FEBS J. **283**, 3232 (2016).

[87]M. F. La Russa and L. S. Qi, Mol. Cell. Biol. **35**, 3800 (2015).

[88]M. Chen, Q. Du, H. Y. Zhang, X. Wang, and Z. Liang, Expert Rev. Mol. Diagn. **7**, 281 (2007).

[89]S. Mockenhaupt, S. Grosse, D. Rupp, R. Bartenschlager, and D. Grimm, Proc. Natl. Acad. Sci. U. S. A. **112**, E4007 (2015).

[90]A. Franceschini, R. Meier, A. Casanova, S. Kreibich, N. Daga, D. Andritschke, S. Dilling, P. Ramö, M. Emmenlauer, A. Kaufmann, R. Conde-Álvarez, S. H. Low, L. Pelkmans, A. Helenius, W. D. Hardt, C. Dehio, and C. Von Mering, Proc. Natl. Acad. Sci. U. S. A. **111**, 4548 (2014).

[91]I. Smith, P. G. Greenside, T. Natoli, D. L. Lahr, D. Wadden, I. Tirosh, R. Narayan, D. E. Root, T. R. Golub, A. Subramanian, and J. G. Doench, PLoS Biol. **15**, e2003213 (2017).

[92]J. Taylor and S. Woodcock, J Biomol Screen. **20**, 1040 (2015).

[93]Y. Takeda, Y. Harada, T. Yoshikawa, and P. Dai, Biosci. Rep. **38**, 1 (2018).

[94]Y. Zhang, W. Li, T. Laurent, and S. Ding, J. Cell Sci. **125**, 5609 (2012).

[95]P. Hou, Y. Li, X. Zhang, C. Liu, J. Guan, H. Li, T. Zhao, J. Ye, W. Yang, K. Liu, J. Ge, J. Xu, Q. Zhang, Y. Zhao, and H. Deng, Science **341**, 651 (2013).

[96]C. Huang, W. Tu, Y. Fu, J. Wang, and X. Xie, Cell Res. **28**, 686 (2018).

[97]W. P. Janzen, Chem. Biol. **21**, 1162 (2014).

[98]S. Chen, Stem Cell Rep. **11**, 1312 (2018).

[99]T. Hart, M. Chandrashekhar, M. Aregger, Z. Steinhart, K. R. Brown, G. MacLeod, M. Mis, M. Zimmermann, A. Fradet-Turcotte, S. Sun, P. Mero, P. Dirks, S. Sidhu, F. P. Roth, O. S. Rissland, D. Durocher, S. Angers, and J. Moffat, Cell **163**, 1515 (2015).

[100]K. Ford, D. McDonald, and P. Mali, J. Mol. Biol. **431**, 48 (2019).

[101]S. Sharma and E. Petsalaki, Int. J. Mol. Sci. **19**, 933 (2018).

[102]O. Shalem, N. E. Sanjana, and F. Zhang, Nat. Rev. Genet. **16**, 299 (2015).

[103]R. J. Ihry, M. R. Salick, D. J. Ho, M. Sondey, S. Kommineni, S. Paula, J. Raymond, B. Henry, E. Frias, Q. Wang, K. A. Worringer, C. Ye, C. Russ, J. S. Reece-Hoyes, R. C. Altshuler, R. Randhawa, Z. Yang, G. McAllister, G. R. Hoffman, R. Dolmetsch, and A. Kaykas, Cell Rep. **27**, 616 (2019).

[104]I. Maza, I. Caspi, A. Zviran, E. Chomsky, Y. Rais, S. Viukov, S. Geula, J. D. Buenrostro, L. Weinberger, V. Krupalnik, S. Hanna, M. Zerbib, J. R. Dutton, W. J. Greenleaf, R. Massarwa, N. Novershtern, and J. H. Hanna, Nat. Biotechnol. **33**, 769 (2015).

[105]D. De, D. Halder, I. Shin, and K. K. Kim, Chem. Soc. Rev. **46**, 6241 (2017).

[106]Y. Buganim, D. A. Faddah, and R. Jaenisch, Nat. Rev. Genet. **14**, 427 (2013).

[107]J. S. Kim, H. W. Choi, S. Choi, H. G. Seo, S. H. Moon, H. M. Chung, and J. T. Do, Stem Cells Dev. **23**, 2637 (2014).

[108]A. Tambe and L. Pachter, BMC Bioinf. **20**(1), 32 (2019).

[109]G. Michlits, M. Hubmann, S. H. Wu, G. Vainorius, E. Budusan, S. Zhuk, T. R. Burkard, M. Novatchkova, M. Aichinger, Y. Lu, J. Reece-Hoyes, R. Nitsch, D. Schramek, D. Hoepfner, and U. Elling, Nat. Methods **14**, 1191 (2017).

[110]A. Dixit, O. Parnas, B. Li, J. Chen, C. P. Fulco, L. Jerby-Arnon, N. D. Marjanovic, D. Dionne, T. Burks, R. Raychowdhury, B. Adamson, T. M. Norman, E. S. Lander, J. S. Weissman, N. Friedman, and A. Regev, Cell **167**, 1853 (2016).

[111]P. Datlinger, A. F. Rendeiro, C. Schmidl, T. Krausgruber, P. Traxler, J. Klughammer, L. C. Schuster, A. Kuchler, D. Alpar, and C. Bock, Nat. Methods **14**, 297 (2017).

[112]D. A. Jaitin, A. Weiner, I. Yofe, D. Lara-Astiaso, H. Keren-Shaul, E. David, T. M. Salame, A. Tanay, A. van Oudenaarden, and I. Amit, Cell **167**, 1883 (2016).

[113]S. Xie, J. Duan, B. Li, P. Zhou, and G. C. Hon, Mol. Cell **66**, 285 (2017).

[114]K. A. Tran, S. J. Pietrzak, N. Z. Zaidan, A. F. Siahpirani, S. G. McCalla, A. S. Zhou, G. Iyer, S. Roy, and R. Sridharan, Cell Rep. **27**, 1726 (2019).

[115]Z. Liu, L. Wang, J. D. Welch, H. Ma, Y. Zhou, H. R. Vaseghi, S. Yu, J. B. Wall, S. Alimohamadi, M. Zheng, C. Yin, W. Shen, J. F. Prins, J. Liu, and L. Qian, Nature **551**, 100 (2017).

[116]M. Francesconi, B. Di Stefano, C. Berenguer, L. De Andrés-Aguayo, M. Plana-Carmona, A. Mendez-Lago, A. Guillaumet-Adkins, G. Rodriguez-Esteban, M. Gut, I. G. Gut, H. Heyn, B. Lehner, and T. Graf, eLife **8**, 1 (2019).

[117]B. Treutlein, Q. Y. Lee, J. G. Camp, M. Mall, W. Koh, S. A. M. Shariati, S. Sim, N. F. Neff, J. M. Skotheim, M. Wernig, and S. R. Quake, Nature **534**, 391 (2016).

[118]B. Hwang, J. H. Lee, and D. Bang, Exp. Mol. Med. **50**, 1 (2018).

[119]Q. H. Nguyen, N. Pervolarakis, K. Nee, and K. Kessenbrock, Front. Cell Dev. Biol. **6**, 1 (2018).

[120]L. Haghverdi, A. T. L. Lun, M. D. Morgan, and J. C. Marioni, Nat. Biotechnol. **36**, 421 (2018).

[121]H. T. N. Tran, K. S. Ang, M. Chevrier, X. Zhang, N. Y. S. Lee, M. Goh, and J. Chen, Genome Biol. **21**, 12 (2020).

[122]T. S. Andrews and M. Hemberg, F1000Research **7**, 1740 (2019).

[123]B. Vieth, S. Parekh, C. Ziegenhain, W. Enard, and I. Hellmann, Nat. Commun. **10**, 1 (2019).

[124]C. S. McGinnis, D. M. Patterson, J. Winkler, D. N. Conrad, M. Y. Hein, V. Srivastava, J. L. Hu, L. M. Murrow, J. S. Weissman, Z. Werb, E. D. Chow, and Z. J. Gartner, Nat. Methods **16**, 619 (2019).

[125]D. Shin, W. Lee, J. H. Lee, and D. Bang, Sci. Adv. **5**, eaav2249 (2019).

[126]J. G. Doench, Nat. Rev. Genet. **19**, 67 (2018).

[127]L. Qian, Y. Huang, C. I. Spencer, A. Foley, V. Vedantham, L. Liu, S. J. Conway, J. D. Fu, and D. Srivastava, Nature **485**, 593 (2012).

[128]S. Yui, L. Azzolin, M. Maimets, M. T. Pedersen, R. P. Fordham, S. L. Hansen, H. L. Larsen, J. Guiu, M. R. P. Alves, C. F. Rundsten, J. V. Johansen, Y. Li, C. D. Madsen, T. Nakamura, M. Watanabe, O. H. Nielsen, P. J. Schweiger, S. Piccolo, and K. B. Jensen, Cell Stem Cell **22**, 35 (2018).

[129]A. Kumar, D. C. Wiltz, and K. J. Grande-Allen, Cardiovasc. Eng. Technol. **4**, 16 (2013).

[130]M. C. Lampi and C. A. Reinhart-King, Sci. Transl. Med. **10**, 475 (2018).

[131]A. Kumar, J. K. Placone, and A. J. Engler, Development **144**, 4261 (2017).

[132]M. G. Ondeck, A. Kumar, J. K. Placone, C. M. Plunkett, B. F. Matte, K. C. Wong, L. Fattet, J. Yang, and A. J. Engler, Proc. Natl. Acad. Sci. **116**, 3502 (2019).

[133]A. Kumar, S. K. Thomas, K. C. Wong, V. Lo Sardo, D. S. Cheah, Y. H. Hou, J. K. Placone, K. P. Tenerelli, W. C. Ferguson, A. Torkamani, E. J. Topol, K. K. Baldwin, and A. J. Engler, Nat. Biomed. Eng. **3**, 137 (2019).

[134]C. Plunkett, A. Kumar, J. Yrastorza, Y. Hou, J. K. Placone, G. Grennan, and A. J. Engler, Adv. Healthc. Mater. **9**, 1901366 (2020).

[135]T. L. Downing, J. Soto, C. Morez, T. Houssin, A. Fritz, F. Yuan, J. Chu, S. Patel, D. V. Schaffer, and S. Li, Nat. Mater. **12**, 1154 (2013).

[136]K. Kulangara, A. F. Adler, H. Wang, M. Chellappan, E. Hammett, R. Yasuda, and K. W. Leong, Biomaterials **35**, 5327 (2014).

[137]J. Sia, P. Yu, D. Srivastava, and S. Li, Biomaterials **103**, 1 (2016).

[138]H. Gerardo, A. Lima, J. Carvalho, J. R. D. Ramos, S. Couceiro, R. D. M. Travasso, R. Pires das Neves, and M. Grãos, Sci. Rep. **9**, 1 (2019).

[139]A. J. Engler, S. Sen, H. L. Sweeney, and D. E. Discher, Cell **126**, 677 (2006).

[140]Y. Li, S. Dal-Pra, M. Mirotsou, T. M. Jayawardena, C. P. Hodgkinson, N. Bursac, and V. J. Dzau, Sci. Rep. **6**, 38815 (2016).

[141]S. Y. Song, J. Yoo, S. Go, J. Hong, H. Su Sohn, J. R. Lee, M. Kang, G. J. Jeong, S. Ryu, S. H. L. Kim, N. S. Hwang, K. Char, and B. S. Kim, Theranostics **9**, 6734 (2019).

[142]J. A. de M. Araújo, M. M. Hilscher, D. Marques-Coelho, D. C. F. Golbert, D. A. Cornelio, S. R. Batistuzzo de Medeiros, R. N. Leão, and M. R. Costa, Front. Cell. Neurosci. **12**, 1 (2018).

[143]G. Chen, X. Xu, L. Zhang, Y. Fu, M. Wang, H. Gu, and X. Xie, Cell Res. **24**, 1121 (2014).

[144]H. Li, M. Collado, A. Villasante, K. Strati, S. Ortega, M. Cãamero, M. A. Blasco, and M. Serrano, Nature **460**, 1136 (2009).

[145]A. Somers, J. C. Jean, C. A. Sommer, A. Omari, C. C. Ford, J. A. Mills, L. Ying, A. G. Sommer, J. M. Jean, B. W. Smith, R. Lafyatis, M. F. Demierre, D. J. Weiss, D. L. French, P. Gadue, G. J. Murphy, G. Mostoslavsky, and D. N. Kotton, Stem Cells **28**, 1728 (2010).

[146]V. Lo Sardo, W. Ferguson, G. A. Erikson, E. J. Topol, K. K. Baldwin, and A. Torkamani, Nat. Biotechnol. **35**, 69 (2017).

[147]G. L. Boulting, E. Kiskinis, G. F. Croft, M. W. Amoroso, D. H. Oakley, B. J. Wainger, D. J. Williams, D. J. Kahler, M. Yamaki, L. Davidow, C. T. Rodolfa, J. T. Dimos, S. Mikkilineni, A. B. MacDermott, C. J. Woolf, C. E. Henderson, H. Wichterle, and K. Eggan, Nat. Biotechnol. **29**, 279 (2011).

[148]S. Protze, S. Khattak, C. Poulet, D. Lindemann, E. M. Tanaka, and U. Ravens, J. Mol. Cell. Cardiol. **53**, 323 (2012).

[149]J. X. Chen, M. Krane, M. A. Deutsch, L. Wang, M. Rav-Acha, S. Gregoire, M. C. Engels, K. Rajarajan, R. Karra, E. D. Abel, J. C. Wu, D. Milan, and S. M. Wu, Circ. Res. **111**, 50 (2012).

[150]P. K. Singhal, S. Sassi, L. Lan, P. Au, S. C. Halvorsen, D. Fukumura, R. K. Jain, and B. Seed, Proc. Natl. Acad. Sci. U. S. A. **113**, 122 (2016).

[151]S. R. Ali, S. Ranjbarvaziri, M. Talkhabi, P. Zhao, A. Subat, A. Hojjat, P. Kamran, A. M. S. Müller, K. S. Volz, Z. Tang, K. Red-Horse, and R. Ardehali, Circ. Res. **115**, 625 (2014).

[152]Y. J. Nam, K. Song, X. Luo, E. Daniel, K. Lambeth, K. West, J. A. Hill, J. M. Di Maio, L. A. Baker, R. Bassel-Duby, and E. N. Olson, Proc. Natl. Acad. Sci. U. S. A. **110**, 5588 (2013).

[153]F. Pistollato, D. Canovas-Jorda, D. Zagoura, and A. Price, J. Vis. Exp. **2017**, 1.

[154]J. Kuijlaars, T. Oyelami, A. Diels, J. Rohrbacher, S. Versweyveld, G. Meneghello, M. Tuefferd, P. Verstraelen, J. R. Detrez, M. Verschuuren, W. H. De Vos, T. Meert, P. J. Peeters, M. Cik, R. Nuydens, B. Brône, and A. Verheyen, Sci. Rep. **6**, 36529 (2016).

[155]A. Schuster, H. Erasimus, S. Fritah, P. V. Nazarov, E. van Dyck, S. P. Niclou, and A. Golebiewska, Trends Biotechnol. **37**, 38 (2019).

[156]H. Yamakawa, N. Muraoka, K. Miyamoto, T. Sadahiro, M. Isomi, S. Haginiwa, H. Kojima, T. Umei, M. Akiyama, Y. Kuishi, J. Kurokawa, T. Furukawa, K. Fukuda, and M. Ieda, Stem Cell Rep. **5**, 1128 (2015).

[157]E. Karamariti, A. Margariti, B. Winkler, X. Wang, X. Hong, D. Baban, J. Ragoussis, Y. Huang, J. D. J. Han, M. M. Wong, C. M. Sag, A. M. Shah, Y. Hu, and Q. Xu, Circ. Res. **112**, 1433 (2013).

[158]B. Xie, D. Sun, Y. Du, J. Jia, S. Sun, J. Xu, Y. Liu, C. Xiang, S. Chen, H. Xie, Q. Wang, G. Li, X. Lyu, H. Shen, S. Li, M. Wu, X. Zhang, Y. Pu, K. Xiang, W. Lai, P. Du, Z. Yuan, C. Li, Y. Shi, S. Lu, and H. Deng, Cell Res. **29**, 696 (2019).

[159]Y. Liu, Y. Xue, S. Ridley, D. Zhang, K. Rezvani, X. D. Fu, and H. Wang, PLoS One **9**, e109621 (2014).

[160]J. Drouin-Ouellet, K. Pircs, R. A. Barker, J. Jakobsson, and M. Parmar, Front. Neurosci. **11**, 530 (2017).

[161]X. Lian, J. Zhang, S. Azarin, K. Zhu, L. B. Hazeltine, X. Bao, C. Hsiao, T. Kamp, and S. P. Palecek, Nat. Protoc. **8**, 162 (2013).

[162]J. V. Terrovitis, R. R. Smith, and E. Marbán, Circ. Res. **106**, 479 (2010).

[163]N. Yang, S. Chanda, S. Marro, Y. H. Ng, J. A. Janas, D. Haag, C. E. Ang, Y. Tang, Q. Flores, M. Mall, O. Wapinski, M. Li, H. Ahlenius, J. L. Rubenstein, H. Y. Chang, A. A. Buylla, T. C. Südhof, and M. Wernig, Nat. Methods **14**, 621 (2017).

[164]Z. Guo, L. Zhang, Z. Wu, Y. Chen, F. Wang, and G. Chen, Cell Stem Cell **14**, 188 (2014).

[165]A. Cieślar-Pobuda, V. Knoflach, M. V. Ringh, J. Stark, W. Likus, K. Siemianowicz, S. Ghavami, A. Hudecki, J. L. Green, and M. J. Łos, Biochim. Biophys. Acta–Mol. Cell Res. **1864**, 1359 (2017).

[166]M. F. Naso, B. Tomkowicz, W. L. Perry, and W. R. Strohl, BioDrugs **31**, 317 (2017).

[167]H. Büning and A. Srivastava, Mol. Ther.–Methods Clin. Dev. **12**, 248 (2019).