



Minireview

Single-Cell Toolkits Opening a New Era for Cell Engineering

Sean Lee^{1,2}, Jireh Kim^{1,2}, and Jong-Eun Park^{1,*}

¹Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Korea, ²These authors contributed equally to this work.

*Correspondence: jp24@kaist.ac.kr

<https://doi.org/10.14348/molcells.2021.0002>

www.molcells.org

Since the introduction of RNA sequencing (RNA-seq) as a high-throughput mRNA expression analysis tool, this procedure has been increasingly implemented to identify cell-level transcriptome changes in a myriad of model systems. However, early methods processed cell samples in bulk, and therefore the unique transcriptomic patterns of individual cells would be lost due to data averaging. Nonetheless, the recent and continuous development of new single-cell RNA sequencing (scRNA-seq) toolkits has enabled researchers to compare transcriptomes at a single-cell resolution, thus facilitating the analysis of individual cellular features and a deeper understanding of cellular functions. Nonetheless, the rapid evolution of high throughput single-cell “omics” tools has created the need for effective hypothesis verification strategies. Particularly, this issue could be addressed by coupling cell engineering techniques with single-cell sequencing. This approach has been successfully employed to gain further insights into disease pathogenesis and the dynamics of differentiation trajectories. Therefore, this review will discuss the current status of cell engineering toolkits and their contributions to single-cell and genome-wide data collection and analyses.

Keywords: cell engineering, CRISPR screening, lineage tracing, single-cell multi-omics

INTRODUCTION

Since the first single-cell transcriptome analysis in 2009, the throughput of single-cell transcriptomic techniques has grown exponentially, allowing for a single study to characterize millions of cells (Svensson et al., 2018). Additionally, single-cell approaches are no longer limited to RNA analyses, but can also be used to characterize DNA and proteins (Lee et al., 2020). This powerful technique has been adopted in many fields of the life sciences and has rapidly expanded our understanding of biological systems. For example, a transcriptome and an open chromatin atlas of the embryonic development process of humans and mice are being created, thus providing new insights into the mechanisms by which gene expression modulates an individual’s developmental process (Cao et al., 2019; 2020; He et al., 2020a; Park et al., 2020; Pijuan-Sala et al., 2020). Moreover, the human immune system’s response to COVID-19 is being currently studied at various levels to gain insights into potential molecular mechanisms that could be targeted to control this disease (Sungnak et al., 2020; Zhang et al., 2020). The Human Cell Atlas is an international collaborative initiative that has provided countless researchers with a platform to produce data cooperatively, as well as to compare and analyze their results and focus their efforts towards a single common goal (Panina et al., 2020; Regev et al., 2017; 2018).

The ever-increasing wealth of single-cell data is deepening our understanding of the structure of the human body. How-

Received 31 December, 2020; revised 4 March, 2021; accepted 11 March, 2021; published online 31 March, 2021

eISSN: 0219-1032

©The Korean Society for Molecular and Cellular Biology. All rights reserved.

©This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>.

ever, to further this understanding, a system that efficiently validates new hypotheses is urgently needed. Interestingly, single-cell techniques can be combined with various cell engineering techniques to provide a platform for efficient hypothesis verification, and several single-cell engineering toolkits are being developed based on this strategy. Therefore, this review will address how cell engineering techniques such as CRISPR screening and lineage tracing are being combined with single-cell techniques to usher in a new era of cell engineering (Table 1).

INCREASING THROUGHPUT OF SINGLE-CELL RNA TECHNIQUES

Early single-cell sequencing approaches depended on amplifying the genetic materials of a single cell to create an RNA-seq library (Tang et al., 2009). Later on, two layers of barcode systems were introduced to increase the throughput of single-cell techniques. Among these, cell barcodes (CBCs) are incorporated during reverse transcription (RT) or template switching steps. This allows for the simultaneous preparation of multiple cells by pooling them after cDNA synthesis, increasing the throughput and efficiency of scRNA-seq library preparation. The second barcoding technique consists of a unique molecular identifier (UMI), which corrects the quantification error caused by polymerase chain reaction (PCR) amplification bias by adding random barcodes during the RT reaction (Kivioja et al., 2011).

In order to separate the cells using primers with unique barcode sequences, a multi-well plate system was first im-

plemented, which allowed for a throughput of 100 to 1,000 cells (Hashimshony et al., 2012; Islam et al., 2011; Ramsköld et al., 2012) (Fig. 1). Afterward, the development of microfluidic systems in which each cell is mixed with CBC-specific RT primer conjugated-beads within a single droplet further increased the throughput to more than 10,000 cells (Klein et al., 2015; Macosko et al., 2015). One of the most recent advancements in this field includes *in situ* combinatorial indexing, which fixes cells to confine the RNA within them, then uses a split-and-pool technique (the cells are sequentially divided into well plates then recombined several times) to attach barcodes and adapter sequences in several combinations. This method increased the throughput of single-cell RNA sequencing (scRNA-seq) from hundreds of thousands to millions of cells (Cao et al., 2017; Rosenberg et al., 2018).

In order to sequence the genetic materials of single cells, solid tissues must first be dissociated. However, it is difficult to dissociate tissues with hard-to-release cell types while avoiding damage to fragile tissues. Many attempts have been made to achieve this balance, such as cryopreservation and methanol fixation. However, each of these approaches has its unique limitations (major loss of epithelial cell types and ambient RNA leakage, respectively), which negatively affect the final scRNA-seq results (Denisenko et al., 2020; Slyper et al., 2020). Single-nucleus RNA-seq (snRNA-seq) was developed to solve these problems. snRNA-seq decouples tissue acquisition from immediate sample processing, thus circumventing the inherent difficulties of obtaining fresh tissues for scRNA-seq analyses, as well as the potential loss of sensitive cells due to enzymatic digestion. Therefore, this strategy can be ap-

Table 1. Representative studies on single-cell engineering

Name of the technique	Perturbation mechanism	Detection mechanism	No. of perturbations	No. of cells	Modularity	Reference
Perturb-seq	CRISPRi	sgRNA barcode	67 sgRNAs (24 genes)	~30,000	RNA	Dixit et al. (2016)
CRISP-seq	CRISPR KO	sgRNA barcode	57 sgRNAs (22 genes)	6,144	RNA	Jaitin et al. (2016)
Mosaic-seq	CRISPRi	sgRNA barcode	241 sgRNAs (71 enhancers)	12,444	RNA	Xie et al. (2017)
CROP-seq	CRISPR KO	sgRNA barcode	48 sgRNAs (20 genes)	N/A	RNA	Datlinger et al. (2017)
Perturb-ATAC-seq	CRISPRi	sgRNA direct capture/ barcode	~190 sgRNAs (63 genes)	~4,300	Chromatin	Rubin et al. (2019)
ECCITE-seq	CRISPR KO	sgRNA direct capture	N/A	N/A	RNA, surface protein	Mimitou et al. (2019)
Convert-seq	cDNA	cDNA sequence	20 genes (transcription factors)	466	RNA	Luginbühl et al. (2019)
Perturb-CITE-seq	CRISPR KO	sgRNA barcode	~750 sgRNAs (~250 genes)	~218,000	Surface protein	Frangieh et al. (2020)
Spear-ATAC-seq	CRISPRi	sgRNA DNA PCR	414 sgRNAs	104,592	Chromatin	Pierce et al. (2020)
Targeted-Perturb-seq (TAP-seq)	CRISPRi	sgRNA barcode	1,790 enhancers	231,667	RNA	Schraivogel et al. (2020)
Perturb-seq	cDNA	cDNA barcode	200 cancer gene variants	>300,000	RNA	Ursu et al. (2020)

Name of technique used, perturbation mechanism, method to detect perturbation identity, number of perturbations tried together, number of single-cells analyzed, modularity of single-cell method, and reference to original paper is denoted.

N/A, not applicable.

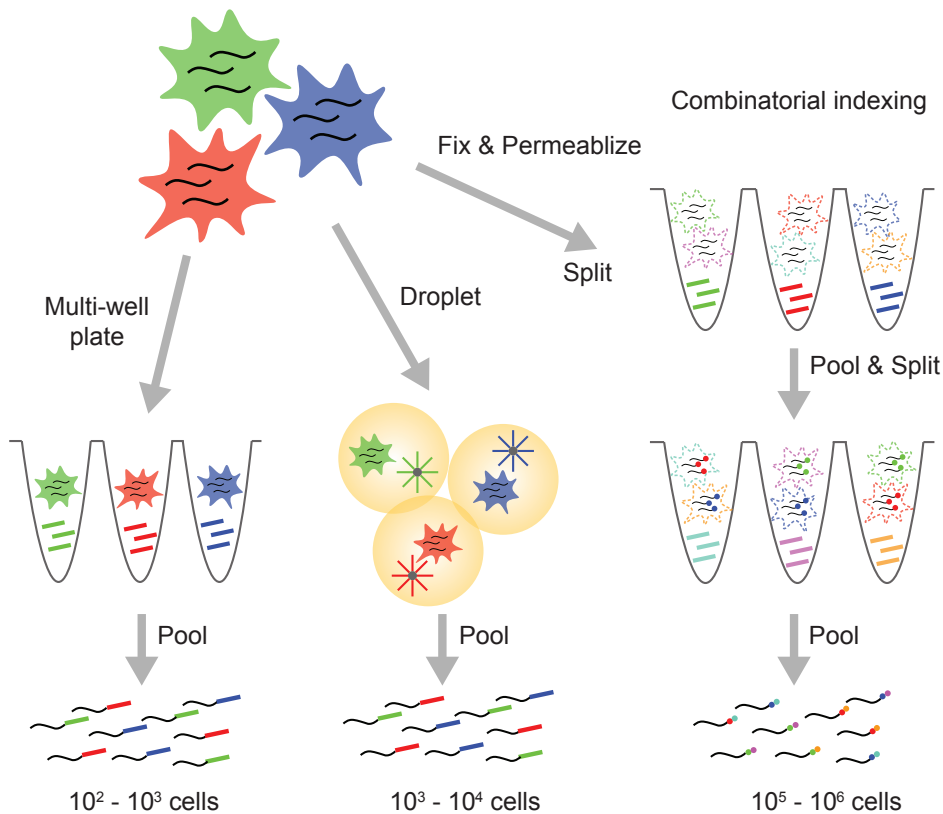


Fig. 1. Development of scRNA-seq techniques with increasing throughput. To attach unique cell barcodes to RNA molecules derived from individual cells, the cells and barcoded primers are isolated together by (1) manually sorting them into microwell plates, (2) generating lipid droplets using microfluidic systems, and (3) repetitive split-and-pooling processes to generate combinatorial barcodes.

plied to hard-to-dissociate tissues, such as tissues rich in cell types (e.g., neurons, adipocytes, and skeletal muscle cells), archived frozen clinical materials, and tissues that must be frozen to register into specific coordinates. Moreover, given that snRNA-seq can be used to handle minute frozen specimens, large-scale studies from tissue atlases to longitudinal clinical trials and human genetics can be performed (Ernst et al., 2020; Gaublotte et al., 2019; Rozenblatt-Rosen et al., 2020).

Cell hashing was recently introduced to further increase the throughput of scRNA-seq. In this technique, each sample is labeled with unique 'hashtag' barcodes using oligonucleotide-conjugated antibodies (Stoeckius et al., 2017; 2018). This can be also applied to snRNA-seq using DNA-barcoded antibodies targeting the nuclear pore complex (Gaublotte et al., 2019). Moreover, MULTI-seq utilizes chemically modified oligos that can directly stain cellular membranes, making the staining process much easier. Similarly, short barcoding oligo (SBO) barcoding introduces oligos into cells using liposomal transfection, thus enabling single-cell experiment multiplexing. sci-Plex is another newly developed strategy that overcomes this issue by directly labeling nuclei with unmodified single-stranded DNA oligos (Srivatsan et al., 2020). Combined with the sci-RNA-seq combinatorial indexing technique, this approach provides a platform to multiplex hundreds of different conditions, thus rendering a total throughput of millions of cells.

The process of demultiplexing and doublet detection was further enhanced with the creation of demuxlet, a compu-

tational algorithm inspired by algorithms that were initially developed to detect DNA contamination in sequencing samples. Even without an oligo or antibody, demuxlet allows researchers to hash their samples even when only the genotype is available. Additionally, this approach can simultaneously demultiplex and detect doublets from more than two individuals from multiplexed Drop-seq using genetic variations, an achievement that was thought to be impossible prior to the development of this algorithm (Kang et al., 2018). The development of SoupCell then allowed for an increased genotype calling efficiency, thus enabling the determination of genotypes *de novo* even without prior genotype data (Heaton et al., 2020). With these developments, the processing and formatting of single-cell research data are becoming more consistent and straightforward, thus progressively increasing the relative ease of data sharing between laboratories, as well as the output from a myriad of engineering platforms.

COUPLING SINGLE-CELL TECHNIQUES WITH CELL ENGINEERING TOOLKITS

In the field of functional genomics, diverse cell engineering tools such as shRNAs or CRISPRs have been used to modify gene expression. For instance, a pooled screening strategy has been designed to efficiently and simultaneously test the function of multiple genes (Sharma and Petsalaki, 2018). In pooled screening, cells are targeted by a pool of viral vector libraries with shRNAs or CRISPR sgRNAs, and the relative

enrichment of specific DNA sequences in cells with certain phenotypes is measured to identify the genes associated with that phenotype (Joung et al., 2017; Sanjana, 2017; Sims et al., 2011). Despite being favorably used to perform efficient and scalable parallel cell modifications, the effectiveness of this pooled screening approach is limited to simple phenotypes only.

To perform pooled screening with scRNA-seq readouts, genetic modifications (e.g., sgRNAs) must be detected with a scRNA-seq technique. However, all available scRNA-seq platforms rely on oligo-dT priming prior to cDNA synthesis, thus capturing only polyadenylated (poly(A)) RNA transcripts. Nonetheless, most genetic modifications used in pooled screening (sgRNAs or shRNAs) do not contain poly(A) tails. In 2016, several independent studies overcame this issue by inserting barcodes associated with individual sgRNAs into poly(A) reporter transcripts, which resulted in successful CRISPR screening at a single-cell resolution (Adamson et al., 2016; Datlinger et al., 2017; Dixit et al., 2016; Jaitin et al., 2016; Xie et al., 2017) (Fig. 2). Further, Jaitin et al. (2016) developed CRISP-seq, an approach that combines pooled CRISPR screening with scRNA-seq. Specifically, CRISPR inter-

ference (CRISPRi) was combined with a poly(A) unique guide index (UGI) to characterize cellular responses at a single-cell resolution. Similarly, Dixit et al. (2016) developed Perturb-seq, a conceptually similar method of enhancing perturbation analysis utilizing droplet-based microfluidics, thus replacing the conventional CRISP-seq micro-well plate-based methods. Coupling mosaic single-cell analysis with indexed CRISPR sequencing (Mosaic-seq) enabled the successful development of a lentiviral dCas9-KRAB-blast vector that contained the epigenetic modifier KRAB, a repressor of enhancer function, to quantify enhancer repression at a single-cell resolution (Xie et al., 2017). Moreover, recent advances have enabled the direct detection of sgRNAs by scRNA-seq by incorporating a sgRNA-specific RT primer (Replogle et al., 2020), thus facilitating CRISPR scRNA-seq screening without the need to build complex barcoded libraries.

Since its development, CRISPR scRNA-seq screening has been applied to address diverse biological questions. For example, Norman et al. (2019) applied Perturb-seq to manipulate a large number of gene pairs and measure the resulting changes in cell state. The authors then created a gene interaction (GI) manifold (high-dimensional surface) that can be

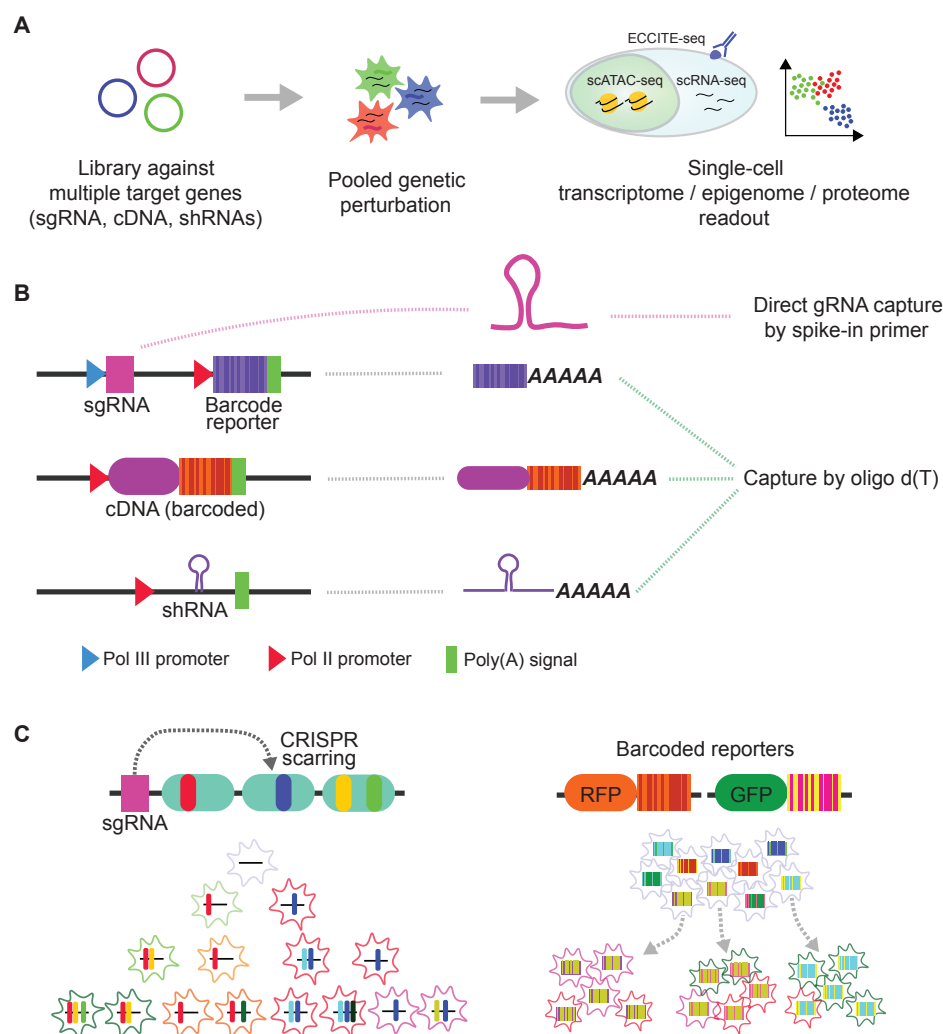


Fig. 2. General schemes and designs of single-cell screening approaches. (A) General scheme of single-cell screening approaches. A library of gene perturbations is introduced into target cells as a pool and their effects are measured at a single-cell resolution to evaluate different molecular endpoints. (B) Design of cell engineering vectors to be detected by single-cell techniques. Gene perturbation targets are identified by incorporating the identifier sequences into transcripts that can be identified by single-cell techniques. (C) Single-cell lineage tracing methods. CRISPR scarring or barcode labeling techniques are used to generate diverse pseudo-genotypes in developing/differentiating cells, which can be detected at the RNA or DNA level in combination with single-cell transcriptome analyses. RFP, red fluorescent protein; GFP, green fluorescent protein.

interpreted and modeled to gain insights into how complex phenotypes emerge. Such large-scale GI analyses may render important insights into how complex, multigenic interactions govern biological traits and disease risks, such as synthetic lethal interactions in cancer and the discovery of gene targets that lessen the severity of genetic diseases (Norman et al., 2019). To cite another example, Jin et al. (2020) applied Perturb-seq *in vivo* to enable the simultaneous assessment of individual phenotypes within a *de novo* likelihood of failure (LoF) risk gene panel. Afterward, the authors identified changes in gene expression and covariation (modularity) of ASD/ND risk gene modules in the developing brain by imposing frameshift mutations in ASD/ND risk genes in mice through *in utero* infection using CRISPR-Cas9 followed by single-cell transcriptomic analysis of perturbed cells from the early postnatal brain (Jin et al., 2020).

Even though Perturb-seq was an instrumental breakthrough in scRNA-seq techniques, its widespread implementation remains limited due to its inherent flaws. Specifically, this approach is prohibitively costly (even for non-genome-scale screens), lowly expressed genes and small effects are not efficiently measured, and a multiple-testing problem greatly undermines data analysis. To solve this problem, Targeted Perturb-seq (TAP-seq) amplifies genes of interest (rather than the whole transcriptome), thus lowering sequencing requirements up to 50-fold. This solves the multiple-hypothesis testing problem encountered in whole transcriptome screens, increases the sensitivity towards small expression changes and lowly expressed genes, and enables the efficient retrieval of sgRNA identities. This decrease in requisites and increase in sensitivity has broadened the applicability of TAP-seq to a wide range of functional genomics applications, including studies where phenotypes of interest are characterized by expression changes in small gene sets (Schraivogel et al., 2020).

In addition to CRISPR-based engineering, other gene-editing tools are also being incorporated into single-cell analyses. For example, Exogenous cDNAs can be identified in scRNA-seq by incorporating specific markers or barcode sequences into their UTR. This scheme has been applied to test the effect of various oncogene variants (Ursu et al., 2020) or to identify transcription factor (TF) sets that can transdifferentiate fibroblasts to neurons (Luginbühl et al., 2019). shRNA-mediated scRNA-seq screens have also been reported, including the identification of shRNAs expressed from pol II transcripts using scRNA-seq (Aarts et al., 2017).

EXPANDING THE APPLICABILITY OF SINGLE-CELL ENGINEERING TOOLKITS

The range of cellular features that can be analyzed with single-cell engineering toolkits evolves as novel screening systems for single-cell sequencing methods are developed. Early screening systems and toolkits were limited to scRNA analyses; however, recent advancements have shifted this focus toward multi-modal single-cell analyses. Various data from different ‘omics’ such as transcriptomics and proteomics (i.e., “multi-omics”) are also being integrated with cell engineering toolkits in single-cell sequencing.

ATAC-seq (assay for transposase-accessible chromatin using sequencing) is an approach that utilizes Tn5 transposases to tag regulatory regions for chromatin accessibility (Buenrostro et al., 2013). Through many collaborative efforts, scATAC-seq was developed to identify chromatin accessibility variations between cell subpopulations within a sample at single-cell resolution (Buenrostro et al., 2015; Chen et al., 2018; Cusanovich et al., 2015). Perturb-ATAC-seq is a method that combines CRISPR guide RNAs and open chromatin sites detected through ATAC-seq with multiplexed CRISPR interference or knockout. Rubin et al. (2019) utilized this approach to compare changes in chromatin states and B lymphoblasts landscapes during CRISPR modifications in broadly-expressed, lineage-specific *trans*-factors. The subsequent mapping of epistatic relationships revealed the role of *trans*-factors in development and disease onset depending on varying combinations of depletion. In the same year, a single-cell gene-editing tool based on scATAC-seq (Spear-ATAC-seq) read-out (i.e., a modified droplet-based scATAC-seq protocol utilizing CRISPRi) was developed. Unlike previous methods, Spear-ATAC-seq relies on reading sgRNA spacer sequences from genomic DNA instead of RNA transcripts. Pierce et al. (2020) reported that this approach could be used to accurately and cost-effectively (\$9.80/cell for ~30 Perturb-ATAC runs compared to one \$0.46/cell for Spear-ATAC) characterize the effects of CRISPR/Cas9-based perturbations on chromatin accessibility.

Much like the development of Perturb-ATAC-seq, Frangieh et al. (2020) recently developed Perturb-CITE-seq, a method that integrates Perturb-seq and CITE-seq to conduct scRNA-seq profiling and epitome sequencing of single-cell surface proteins under specific perturbations caused by multiplexed CRISPR mediated gene inactivation. Cellular Indexing of Transcriptomes and Epitopes (CITE-seq) is a high-throughput technique that is widely used to quantify single-cell mRNA and surface protein expression through oligonucleotide-labeled antibodies (Stoeckius et al., 2017). The integration of single-cell CRISPR-Cas9 functional screening and CITE-seq allows for more efficient identification of genes, as many relevant phenotypes are known to be best understood functionally at the protein level rather than at the transcript level (Yang et al., 2020). Expanded CRISPR-compatible CITE by sequencing (ECCITE-seq) was developed by Mimitou et al. (2019) to improve the original CITE-seq toolkit even further to reach a new level of multi-modal applicability in single-cell multi-omics research. ECCITE-seq can be used to characterize not only the transcriptome, but also the cell hashing, T cell antigen receptors (clonotypes), CRISPR perturbations, and surface proteins with sgRNA of single cells.

Until recently, only surface proteins could be analyzed at a single-cell resolution due to the intricacies of accessing the cell interior for intracellular protein sequencing via cellular fixation (Saliba et al., 2014). Intracellular staining and sequencing (INs-seq) was developed by Katzenelenbogen et al. (2020) to enable intracellular protein immunodetection via cellular fixation, which could then be analyzed using scRNA-seq. Using this approach, the authors fixed and permeabilized cells using a fixative based on methanol and ammonium sulfate solutions that precipitates proteins, inhibits enzymatic activity, and preserves RNA. More importantly, this approach

enables immuno-intracellular staining while preserving mRNA integrity. Permeabilized cells were then intracellularly labeled with fluorophore-conjugated antibodies, then sorted by fluorescence-activated cell sorting (FACS) followed by scRNA-seq utilizing plate-based or microfluidics-based approaches. Comparisons with commonly used paraformaldehyde (Thomsen et al., 2016), methanol (Alles et al., 2017), and dithiobis(succinimidyl propionate) (Attar et al., 2018) fixation methods demonstrated that INs-seq preserved mRNA more effectively. INs-seq can therefore be used to characterize intracellular and post-translationally-modified proteins (PTM), signaling pathways, TF, and metabolism-related proteins at a single-cell resolution when coupled with scRNA-seq, thus enabling the analysis of intracellular signals that may not be typical of specific cell lineages.

LINEAGE TRACING COUPLED WITH SINGLE-CELL RNA SEQUENCING

Deriving lineage relationships between cells within a developing organism has long been a primary focus in the field of developmental biology, with fate mapping methods constantly being created to achieve this goal. Bulk lineage tracing has existed long before single-cell sequencing. However, the possibility of internal heterogeneity in specific cell populations or a few off-target cells has not been comprehensively assessed (Wagner and Klein, 2020). Single-cell screening systems have evolved to not only identify single-cell features, thus increasing the precision of the labeling process to limit potential cellular heterogeneity, but also to link the data obtained from various cells within a cell type to elucidate potential relationships. Lineage tracing can therefore provide critical insights into the pluripotency of stem cells. In turn, the findings gathered from such experiments could be used to treat diseases and aid in the development of regenerative medicine. Retroviral labeling is an earlier method of lineage tracing that utilizes libraries containing reporter transgenes (e.g., beta-galactosidase and GFP) and DNA fragment barcodes for clonal relationship analysis via PCR amplification followed by sequencing (Walsh and Cepko, 1992). However, although this procedure has been successfully used in the past to elucidate lineage relationships, it has a few limitations that must be considered prior to its implementation. Concretely, this method cannot be used to characterize cell lines at a single-cell resolution and is therefore unsuitable in instances where higher-dimensional and more complex data is needed, thus highlighting the limitations of this approach compared to more recent techniques. Retroviral vectors can become spontaneously silent during the experiment, and therefore certain experimental paradigms such as histochemical labeling may become more challenging (Ginsberg and Che, 2004; Mayer et al., 2015). Furthermore, with the barcode method implemented in retroviral labeling, only the cells with the ability to divide can pass down the barcode to their progeny, thus limiting the cells and methods that can be used.

Recently developed methods have begun to address many of the limitations of previous techniques such as retroviral labeling, including the implementation of microfluidics platforms in combination with scRNA-seq to characterize

both interclonal and intraclonal variability of CD8 T-cells (Kimmerling et al., 2016). One of the most significant recent advancements in this technique was reported by McKenna et al. (2016), in which the authors developed the Genome Editing of Synthetic Target Arrays for Lineage Tracing (GESTALT) method. This approach utilizes CRISPR-Cas9 genome-editing technology to create unique mutation barcode patterns (McKenna et al., 2016). These unique mutations are accumulated after several sequential cellular divisions, which are then recovered using targeted sequencing, thus enabling the identification of lineage relationships between cells. In turn, this allows for a more efficient analysis of development (Kalhor et al., 2018) and does not have as many limitations as previous methods, including potential silencing from retroviral labeling. A disadvantage of GESTALT, however, is that it cannot yet determine the precise anatomical position and cell type of each assayed cell, meaning that other methods may be needed if the aforementioned information is required in a specific experiment. Raj et al. (2018) recently developed scGESTALT, which makes use of large-scale transcriptional profiling via the inDrops microfluidic platform (Zilionis et al., 2017) to extend the CRISPR-Cas9 lineage tracing abilities of the original GESTALT technique (Raj et al., 2018). Therefore, this approach offers a means for more efficient cell fate mapping and the elucidation of clonal relationships at the single-cell level.

ScarTrace is another technique that utilizes CRISPR-Cas9 technology to induce double-strand breaks, thus resulting in different-length insertions or deletions at various positions (scars) to create heritable labels (Alemany et al., 2018). Recent developments have furthered the evolution of CRISPR scarring methods such as ScarTrace and other barcoding methods. For example, iTracer (He et al., 2020b) is a novel technique that takes advantage of previously developed complex barcode libraries (Weinreb et al., 2020) and an inducible Cas9 scarring system similar to the one used in the lineage tracing by nuclease-activated editing of ubiquitous sequences (LINNAEUS) method (Spanjaard et al., 2018) to reconstruct the lineage trajectories and fate decisions of induced pluripotent stem cells (iPSCs). CellTagging is another recently developed method that allows for the parallel capture of lineage information and cell identity through a combinatorial cell indexing approach with lentiviral barcoding (Kong et al., 2020). This protocol consists of generating complex plasmid and lentivirus CellTag libraries to label cells, followed by sequential CellTagging throughout a given biological process. Lentiviral barcodes were incorporated via the integration of a constitutively expressed GFP-encoding gene with random barcodes engineered into the 3' untranslated region sequence, after which additional barcoding rounds were applied to mark successive lineage restriction events, followed by scRNA-seq analysis. Biddy et al. (2018) used CellTagging on mouse fibroblast cells, focusing on the direct reprogramming of fibroblasts into induced endoderm progenitors (iEPs). Therefore, increasing the throughput of novel lineage tracing methods could provide a robust and more streamlined basis for the elucidation of relationships between higher populations of cells within a cell line.

CONCLUSION

This review discussed not only the recent developments in scRNA-seq technology and its advantages, but also the increased applicability of single-cell techniques, particularly when coupled with cell engineering toolkits. The combination of genetic perturbation approaches (e.g., CRISPR, CRISPRi, CRISPRa, TF over-expression, and lineage tracing, among others) with multimodal, single-cell, genome-wide phenotyping is revolutionizing the field of functional genomics by generating biometric data at unprecedented rates. Medical research benefits greatly from single-cell modification technologies, as various responses to a myriad of TFs can be studied at a single-cell resolution across multiple cell lines. However, despite the recent developments of perturbation-mediated toolkits such as Perturb-ATAC-seq and Perturb-CITE-seq, many techniques are yet to implement these perturbation methods. Therefore, the future co-implementation of genetic perturbation and lineage tracing into novel methods will be essential to further the field of single-cell multimodal omics. These changes will significantly improve the depth to which the clonal relationships between cells can be characterized and aid in the development of stem cell research. Moreover, INS-seq's ability to analyze intracellular proteins is one of the most recent breakthroughs in the effort to dissect the role of individual proteins in single-cell physiology. Further integration of intracellular protein analysis and continuous expansion into new modalities will likely enhance our understanding of the various interacting cell lines that comprise the human body, as well as the potential molecular and cellular mechanisms that may be targeted to treat and prevent diseases. Finally, increasing the throughput of current approaches from hundreds of thousands to millions of cells would allow for a significantly more streamlined whole-genome profiling of entire organisms, thus surpassing the capacity of current next-generation sequencing technologies.

ACKNOWLEDGMENTS

This work was supported by the POSCO Science Fellowship of the POSCO TJ Park Foundation.

AUTHOR CONTRIBUTIONS

S.L., J.K., and J.E.P. wrote the manuscript. J.E.P. created the figures.

CONFLICT OF INTEREST

The authors have no potential conflicts of interest to disclose.

ORCID

Sean Lee <https://orcid.org/0000-0003-4892-247X>
Jireh Kim <https://orcid.org/0000-0002-3547-7445>
Jong-Eun Park <https://orcid.org/0000-0002-1687-2423>

REFERENCES

Aarts, M., Georgilis, A., Beniazza, M., Beolchi, P., Banito, A., Carroll, T., Kulisic, M., Kaemena, D.F., Dharmalingam, G., Martin, N., et al. (2017). Coupling shRNA screens with single-cell RNA-seq identifies a dual role for mTOR in reprogramming-induced senescence. *Genes Dev.* *31*, 2085-2098.
Adamson, B., Norman, T.M., Jost, M., Cho, M.Y., Nuñez, J.K., Chen, Y.,

Villalta, J.E., Gilbert, L.A., Horlbeck, M.A., Hein, M.Y., et al. (2016). A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* *167*, 1867-1882.e21.

Aleman, A., Florescu, M., Baron, C.S., Peterson-Maduro, J., and van Oudenaarden, A. (2018). Whole-organism clone tracing using single-cell sequencing. *Nature* *556*, 108-112.

Alles, J., Karaiskos, N., Praktiknjo, S.D., Grosswendt, S., Wahle, P., Ruffault, P.L., Ayoub, S., Schreyer, L., Boltengagen, A., Birchmeier, C., et al. (2017). Cell fixation and preservation for droplet-based single-cell transcriptomics. *BMC Biol.* *15*, 44.

Attar, M., Sharma, E., Li, S., Bryer, C., Cubitt, L., Broxholme, J., Lockstone, H., Kinchen, J., Simmons, A., Piazza, P., et al. (2018). A practical solution for preserving single cells for RNA sequencing. *Sci. Rep.* *8*, 2151.

Biddy, B.A., Kong, W., Kamimoto, K., Guo, C., Wayne, S.E., Sun, T., and Morris, S.A. (2018). Single-cell mapping of lineage and identity in direct reprogramming. *Nature* *564*, 219-224.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* *10*, 1213-1218.

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* *523*, 486-490.

Cao, J., O'Day, D.R., Pliner, H.A., Kingsley, P.D., Deng, M., Daza, R.M., Zager, M.A., Aldinger, K.A., Blecher-Gonen, R., Zhang, F., et al. (2020). A human cell atlas of fetal gene expression. *Science* *370*, eaba7721.

Cao, J., Packer, J.S., Ramani, V., Cusanovich, D.A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S.N., Steemers, F.J., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* *357*, 661-667.

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature* *566*, 496-502.

Chen, X., Miragaia, R.J., Natarajan, K.N., and Teichmann, S.A. (2018). A rapid and robust method for single cell chromatin accessibility profiling. *Nat. Commun.* *9*, 5345.

Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C., and Shendure, J. (2015). Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* *348*, 910-914.

Datlinger, P., Rendeiro, A.F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., Schuster, L.C., Kuchler, A., Alpar, D., and Bock, C. (2017). Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* *14*, 297-301.

Denisenko, E., Guo, B.B., Jones, M., Hou, R., de Kock, L., Lassmann, T., Poppe, D., Clément, O., Simmons, R.K., Lister, R., et al. (2020). Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biol.* *21*, 130.

Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C.P., Jerby-Aron, L., Marjanovic, N.D., Dionne, D., Burks, T., Raychowdhury, R., et al. (2016). Perturb-seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* *167*, 1853-1866.e17.

Ernst, K.J., Okonechnikov, K., Bageritz, J., Mallm, J.P., Wittmann, A., Maaß, K.K., Leible, S., Boutros, M., Pfister, S.M., Zuckermann, M., et al. (2020). Establishment of a simplified preparation method for single-nucleus RNA-sequencing and its application to long-term frozen tumor tissues. *bioRxiv*, <https://doi.org/10.1101/2020.10.23.351809>

Frangieh, C.J., Melms, J.C., Thakore, P.I., Geiger-Schuller, K.R., Ho, P., Luoma, A.M., Cleary, B., Malu, S., Cuoco, M.S., Zhao, M., et al. (2020). Multi-modal pooled Perturb-CITE-Seq screens in patient models define novel mechanisms of cancer immune evasion. *bioRxiv*, <https://doi.org/10.1101/2020.10.23.351809>

org/10.1101/2020.09.01.267211

Gaublomme, J.T., Li, B., McCabe, C., Knecht, A., Yang, Y., Drokhllyansky, E., Van Wittenberghe, N., Waldman, J., Dionne, D., Nguyen, L., et al. (2019). Nuclei multiplexing with barcoded antibodies for single-nucleus genomics. *Nat. Commun.* *10*, 2907.

Ginsberg, S.D. and Che, S. (2004). Combined histochemical staining, RNA amplification, regional, and single cell cDNA analysis within the hippocampus. *Lab. Invest.* *84*, 952-962.

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.

He, P., Williams, B.A., Trout, D., Marinov, G.K., Amrhein, H., Berghella, L., Goh, S.T., Plajzer-Frick, I., Afzal, V., Pennacchio, L.A., et al. (2020a). The changing mouse embryo transcriptome at whole tissue and single-cell resolution. *Nature* *583*, 760-767.

He, Z., Gerber, T., Maynard, A., Jain, A., Petri, R., Santel, M., Ly, K., Sidow, L., Sanchís-Calleja, F., Riesenberger, S., et al. (2020b). Lineage recording reveals dynamics of cerebral organoid regionalization. *bioRxiv*, <https://doi.org/10.1101/2020.06.19.162032>

Heaton, H., Talman, A.M., Knights, A., Imaz, M., Gaffney, D.J., Durbin, R., Hemberg, M., and Lawnczak, M.K.N. (2020). Souporecell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat. Methods* *17*, 615-620.

Islam, S., Kjällquist, U., Moliner, A., Zajac, P., Fan, J.B., Lönnerberg, P., and Linnarsson, S. (2011). Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* *21*, 1160-1167.

Jaitin, D.A., Weiner, A., Yofe, I., Lara-Astiaso, D., Keren-Shaul, H., David, E., Salame, T.M., Tanay, A., van Oudenaarden, A., and Amit, I. (2016). Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell* *167*, 1883-1896.e15.

Jin, X., Simmons, S.K., Guo, A., Shetty, A.S., Ko, M., Nguyen, L., Jokhi, V., Robinson, E., Oylar, P., Curry, N., et al. (2020). In vivo Perturb-Seq reveals neuronal and glial abnormalities associated with autism risk genes. *Science* *370*, eaaz6063.

Joung, J., Konermann, S., Gootenberg, J.S., Abudayyeh, O.O., Platt, R.J., Brigham, M.D., Sanjana, N.E., and Zhang, F. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* *12*, 828-863.

Kalhor, R., Kalhor, K., Mejia, L., Leeper, K., Graveline, A., Mali, P., and Church, G.M. (2018). Developmental barcoding of whole mouse via homing CRISPR. *Science* *361*, eaat9804.

Kang, H.M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., McCarthy, E., Wan, E., Wong, S., Byrnes, L., Lanata, C.M., et al. (2018). Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* *36*, 89-94.

Katzenelenbogen, Y., Sheban, F., Yalin, A., Yofe, I., Svetlichnyy, D., Jaitin, D.A., Bornstein, C., Moshe, A., Keren-Shaul, H., Cohen, M., et al. (2020). Coupled scRNA-seq and intracellular protein activity reveal an immunosuppressive role of TREM2 in cancer. *Cell* *182*, 872-885.e19.

Kimmerling, R.J., Lee Szeto, G., Li, J.W., Genshaft, A.S., Kazer, S.W., Payer, K.R., de Riba Borrajo, J., Blainey, P.C., Irvine, D.J., Shalek, A.K., et al. (2016). A microfluidic platform enabling single-cell RNA-seq of multigenerational lineages. *Nat. Commun.* *7*, 10220.

Kivioja, T., Vähärautio, A., Karlsson, K., Bonke, M., Enge, M., Linnarsson, S., and Taipale, J. (2011). Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* *9*, 72-74.

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* *161*, 1187-1201.

Kong, W., Biddy, B.A., Kamimoto, K., Amrute, J.M., Butka, E.G., and Morris, S.A. (2020). CellTagging: combinatorial indexing to simultaneously map

lineage and identity at single-cell resolution. *Nat. Protoc.* *15*, 750-772.

Lee, J., Hyeon, D.Y., and Hwang, D. (2020). Single-cell multiomics: technologies and data analysis methods. *Exp. Mol. Med.* *52*, 1428-1442.

Luginbühl, J., Kouno, T., Nakano, R., Chater, T.E., Sivaraman, D.M., Kishima, M., Roudnický, F., Carninci, P., Plessy, C., and Shin, J.W. (2019). Decoding neuronal diversity by single-cell Convert-seq. *bioRxiv*, <https://doi.org/10.1101/600239>

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* *161*, 1202-1214.

Mayer, C., Jaglin, X.H., Cobbs, L.V., Bandler, R.C., Streicher, C., Cepko, C.L., Hippenmeyer, S., and Fishell, G. (2015). Clonally related forebrain interneurons disperse broadly across both functional areas and structural boundaries. *Neuron* *87*, 989-998.

McKenna, A., Findlay, G.M., Gagnon, J.A., Horwitz, M.S., Schier, A.F., and Shendure, J. (2016). Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* *353*, aaf7907.

Mimitou, E.P., Cheng, A., Montalbano, A., Hao, S., Stoeckius, M., Legut, M., Roush, T., Herrera, A., Papalexis, E., Ouyang, Z., et al. (2019). Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* *16*, 409-412.

Norman, T.M., Horlbeck, M.A., Replogle, J.M., Ge, A.Y., Xu, A., Jost, M., Gilbert, L.A., and Weissman, J.S. (2019). Exploring genetic interaction manifolds constructed from rich single-cell phenotypes. *Science* *365*, 786-793.

Panina, Y., Karagiannis, P., Kurtz, A., Stacey, G.N., and Fujibuchi, W. (2020). Human Cell Atlas and cell-type authentication for regenerative medicine. *Exp. Mol. Med.* *52*, 1443-1451.

Park, J.E., Jardine, L., Gottgens, B., Teichmann, S.A., and Haniffa, M. (2020). Prenatal development of human immunity. *Science* *368*, 600-603.

Pierce, S.E., Granja, J.M., and Greenleaf, W.J. (2020). High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. *bioRxiv*, <https://doi.org/10.1101/2020.11.02.364265>

Pijuan-Sala, B., Wilson, N.K., Xia, J., Hou, X., Hannah, R.L., Kinston, S., Calero-Nieto, F.J., Poirion, O., Preissl, S., Liu, F., et al. (2020). Single-cell chromatin accessibility maps reveal regulatory programs driving early mouse organogenesis. *Nat. Cell Biol.* *22*, 487-497.

Raj, B., Gagnon, J.A., and Schier, A.F. (2018). Large-scale reconstruction of cell lineages using single-cell readout of transcriptomes and CRISPR-Cas9 barcodes by scGESTALT. *Nat. Protoc.* *13*, 2685-2713.

Ramsköld, D., Luo, S., Wang, Y.C., Li, R., Deng, Q., Faridani, O.R., Daniels, G.A., Khrebtkova, I., Loring, J.F., Laurent, L.C., et al. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* *30*, 777-782.

Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al.; Human Cell Atlas Meeting Participants (2017). The Human Cell Atlas. *Elife*, *6*, e27041.

Regev, A., Teichmann, S., Rozenblatt-Rosen, O., Stubbington, M., Ardlie, K., Amit, I., Arlotta, P., Bader, G., Benoist, C., Biton, M., et al.; Human Cell Atlas Organizing Committee (2018). *arXiv*, <http://arxiv.org/abs/1810.05192>

Replogle, J.M., Norman, T.M., Xu, A., Hussmann, J.A., Chen, J., Cogan, J.Z., Meer, E.J., Terry, J.M., Riordan, D.P., Srinivas, N., et al. (2020). Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* *38*, 954-961.

Rosenberg, A.B., Roco, C.M., Muscat, R.A., Kuchina, A., Sample, P., Yao, Z., Grayback, L.T., Peeler, D.J., Mukherjee, S., Chen, W., et al. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* *360*, 176-182.

Rozenblatt-Rosen, O., Regev, A., Oberdoerffer, P., Nawy, T., Hupalowska, A.,

- Rood, J.E., Ashenberg, O., Cerami, E., Coffey, R.J., Demir, E., et al.; Human Tumor Atlas Network (2020). The Human Tumor Atlas Network: charting tumor transitions across space and time at single-cell resolution. *Cell* **181**, 236-249.
- Rubin, A.J., Parker, K.R., Satpathy, A.T., Qi, Y., Wu, B., Ong, A.J., Mumbach, M.R., Ji, A.L., Kim, D.S., Cho, S.W., et al. (2019). Coupled single-cell CRISPR screening and epigenomic profiling reveals causal gene regulatory networks. *Cell* **176**, 361-376.e17.
- Saliba, A.E., Westermann, A.J., Gorski, S.A., and Vogel, J. (2014). Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* **42**, 8845-8860.
- Sanjana, N.E. (2017). Genome-scale CRISPR pooled screens. *Anal. Biochem.* **532**, 95-99.
- Schraivogel, D., Gschwind, A.R., Milbank, J.H., Leonce, D.R., Jakob, P., Mathur, L., Korbel, J.O., Merten, C.A., Velten, L., and Steinmetz, L.M. (2020). Targeted Perturb-seq enables genome-scale genetic screens in single cells. *Nat. Methods* **17**, 629-635.
- Sharma, S. and Petsalaki, E. (2018). Application of CRISPR-Cas9 based genome-wide screening approaches to study cellular signalling mechanisms. *Int. J. Mol. Sci.* **19**, 933.
- Sims, D., Mendes-Pereira, A.M., Frankum, J., Burgess, D., Cerone, M.A., Lombardelli, C., Mitsopoulos, C., Hakas, J., Murugaesu, N., Isacke, C.M., et al. (2011). High-throughput RNA interference screening using pooled shRNA libraries and next generation sequencing. *Genome Biol.* **12**, R104.
- Slyper, M., Porter, C.B.M., Ashenberg, O., Waldman, J., Drokhylyansky, E., Wakiro, I., Smillie, C., Smith-Rosario, G., Wu, J., Dionne, D., et al. (2020). A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat. Med.* **26**, 792-802.
- Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N., and Junker, J.P. (2018). Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat. Biotechnol.* **36**, 469-473.
- Strivatsan, S.R., McFaline-Figueroa, J.L., Ramani, V., Saunders, L., Cao, J., Packer, J., Pliner, H.A., Jackson, D.L., Daza, R.M., Christiansen, L., et al. (2020). Massively multiplex chemical transcriptomics at single-cell resolution. *Science* **367**, 45-51.
- Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865-868.
- Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B.Z., Mauck, W.M., 3rd, Smibert, P., and Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224.
- Sungnak, W., Huang, N., Bécavin, C., Berg, M.; HCA Lung Biological Network (2020). SARS-CoV-2 entry genes are most highly expressed in nasal goblet and ciliated cells within human airways. arXiv, <https://www.ncbi.nlm.nih.gov/pubmed/32550242>
- Svensson, V., Vento-Tormo, R., and Teichmann, S.A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.* **13**, 599-604.
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377-382.
- Thomsen, E.R., Mich, J.K., Yao, Z., Hodge, R.D., Doyle, A.M., Jang, S., Shehata, S.I., Nelson, A.M., Shapovalova, N.V., Levi, B.P., et al. (2016). Fixed single-cell transcriptomic characterization of human radial glial diversity. *Nat. Methods* **13**, 87-93.
- Ursu, O., Neal, J.T., Shea, E., Thakore, P.I., Jerby-Arnon, L., Nguyen, L., Dionne, D., Diaz, C., Bauman, J., Mosaad, M.M., et al. (2020). Massively parallel phenotyping of variant impact in cancer with Perturb-seq reveals a shift in the spectrum of cell states induced by somatic mutations. bioRxiv, <https://doi.org/10.1101/2020.11.16.383307>
- Wagner, D.E. and Klein, A.M. (2020). Lineage tracing meets single-cell omics: opportunities and challenges. *Nat. Rev. Genet.* **21**, 410-427.
- Walsh, C. and Cepko, C.L. (1992). Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. *Science* **255**, 434-440.
- Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F.D., and Klein, A.M. (2020). Lineage tracing on transcriptional landscapes links state to fate during differentiation. *Science* **367**, eaaw3381.
- 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.
- Yang, L., Zhu, Y., Yu, H., Cheng, X., Chen, S., Chu, Y., Huang, H., Zhang, J., and Li, W. (2020). scMAGeCK links genotypes with multiple phenotypes in single-cell CRISPR screens. *Genome Biol.* **21**, 19.
- Zhang, J.Y., Wang, X.M., Xing, X., Xu, Z., Zhang, C., Song, J.W., Fan, X., Xia, P., Fu, J.L., Wang, S.Y., et al. (2020). Single-cell landscape of immunological responses in patients with COVID-19. *Nat. Immunol.* **21**, 1107-1118.
- Zilionis, R., Nainys, J., Veres, A., Savova, V., Zemmour, D., Klein, A.M., and Mazutis, L. (2017). Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* **12**, 44-73.