ETH zürich

ETH Library

Single-cell genomics to guide human stem cell and tissue engineering

Review Article

Author(s):
Camp, J. Gray; Wollny, Damian; Treutlein, Barbara

Publication date:
2018

Permanent link:
https://doi.org/10.3929/ethz-b-000459746

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
Nature Methods 15(9), https://doi.org/10.1038/s41592-018-0113-0

Funding acknowledgement:
758877 - Reconstructing human cortex development and malformation with single-cell transcriptomics (EC)

This page was generated automatically upon download from the ETH Zurich Research Collection.
For more information, please consult the Terms of use.

Single-cell genomics guide human stem cell and tissue engineering

J. Gray Camp[1], Damian Wollny[1], Barbara Treutlein[1,2,3]

Affiliations:
[1]Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.
[2]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany.
[3]Technical University Munich, Department of Biosciences, Freising, Germany

Correspondence: J. Gray Camp (gray_camp@eva.mpg.de), Barbara Treutlein (Barbara_treutlein@eva.mpg.de)

Abstract
To understand human development and disease, as well as regenerate damaged tissues, scientists are working to engineer cell types *in vitro* and to create 3D microenvironments in which cells behave physiologically. Single-cell genomic (SCG) technologies are being applied to primary human organs and engineered cells and tissues to generate atlases of cell diversity in these systems at unparalleled resolution. Moving beyond atlases, SCG methods are powerful tools for gaining insight into the engineering and disease process. Here we discuss how single-cell sequencing can be used to optimize human cell and tissue engineering by measuring precision, detecting inefficiencies, and assessing accuracy. We also provide a perspective for how emerging SCG methods can be used to reverse engineer human cells and tissues, and unravel disease mechanisms.

2D and 3D strategies to engineer human cell types

Single-cell genomics (SCG) has recently emerged as a field, owing to the appearance of an exciting set of tools to explore cellular complexity in cell cultures and tissues and to reconstruct differentiation processes[1,2]. In parallel, there has been enormous progress in engineering human cell types and tissues from pluripotent stem cells (PSCs) in culture. In this Perspective, we argue that SCG approaches offer new opportunities to understand existing differentiation protocols and their limitations. These inferences can guide strategies to reverse engineer human organs and explore disease mechanisms.

PSCs grown in two-dimensional culture can be induced to become specific cell types through a regimen of developmental signalling cues, or through the forced expression of cell-type defining transcription factors (TFs)[3,4]. In addition, differentiated somatic cells can be directly converted to another lineage through the expression of TFs or microRNAs, or the application of small molecules[5-7]. Many cell types generated by 2D approaches have been used to understand mechanisms controlling differentiation and barriers to plasticity; they have been used to model disease, and even to regenerate damaged tissues[8-10].

Of course, human tissues are three-dimensional. They are composed of many different cell types that signal each other and coordinate functions at the tissue and organ level. Human PSCs have been shown to self-organize into complex 3D structures, so-called 'organoids', that can recapitulate the morphology and some functionality of tissues including the eye[11], brain[12], liver[13], stomach[14], intestine[15], kidney[16], and others. Organoid technologies offer great promise to model human development and disease, and scientists are starting to assemble modular developing units of tissue (such as the ventral and dorsal telencephalon[17,18]) to create controlled inter-tissue interactions. The ultimate goal is to interconnect multiple organs into systems to study human physiology.

However, there are many bottlenecks and challenges in human cell and tissue engineering. It is often unclear how similar the engineered cell or tissue is to the *in vivo* counterpart, and for many cell and tissue types, existing protocols are inefficient or no protocol exists yet. Even in organoids, it is clear that many tissue-resident cells are not present or sustained in the 3D microenvironment. In the following sections, we discuss how single-cell transcriptomics can address these shortcomings.

Single-cell atlases are optimum references for engineering

It is incredible that only 9 years ago, it was first shown that the transcriptome could be sequenced from a single cell[19]. Since that time, high-throughput single-cell RNA-sequencing (scRNA-

seq) technologies have been used to generate cell atlases from many mouse[20-24] and human tissues[25-28]. These efforts have paved the way for the Human Cell Atlas, a major initiative to systematically catalogue the molecular profiles of each cell type in every human organ and tissue, at multiple developmental time points and across different individuals[29]. These reference maps may serve as a basis for understanding human health, and diagnosing, monitoring, and treating disease. Atlas efforts can also serve as a reference for the cell and tissue engineering fields, and allow quantitative comparisons between engineered and "real" cells (Fig. 1a).

Conventional strategies to assess how well engineered cells and tissues recapitulate primary human counterparts have generally focused on immunohistochemistry, morphology, cell behavior, and other functional readouts. Transcriptome measurements from bulk samples have also helped benchmark the engineering process. However, scRNA-seq can provide an additional quantitative framework to assess the accuracy, precision, and efficiency of cell and organ engineering. First, scRNA-seq can be used to deconstruct the cellular composition of the engineered cells or tissue and the identified cell states can then be compared to the primary counterparts from the reference atlas (Fig. 1b). One can quantitate the accuracy (fraction of *in vitro* transcriptome that resembles that of the corresponding primary cell), precision (ratio of target to off-target lineages in engineered cell culture), and efficiency (proportion of engineered cells that productively advance along the target lineage path) of the engineering process. Furthermore, one can ascertain which cell states are missing from an *in vitro* tissue. Finally, one can reconstruct developmental trajectories and assess whether the path that cells take *in vitro* is the same as the one they take *in vivo*.

There are a number of challenges in comparing reference and engineered systems. Ideally, all data should be generated using the same chemistry and technical platform. However, diverse methods exist to generate single-cell transcriptomes, which can confound quantitative comparisons due to technical differences. Computational approaches are being developed that allow data sets generated from different methods or species to be integrated[30-32], and that can project clusters from one dataset onto another[33]. Even with the same platform, there can be technical differences related to the dissociation of primary tissues versus cultured cells, including higher concentrations of dissociation enzymes or mechanical stress resulting from tissue perfusion or trituration. Optimization of tissue dissociation is therefore a critical aspect of any single-cell genomic experiment, and the discovery[34] and correction of dissociation-induced artifacts should be explored during data analysis. It can also be unclear how the temporal dynamics of engineered and reference cells correspond. Aligning differentiation trajectories can enable these comparisons[35,36] (Fig. 1c). Finally, an optimum reference atlas would ideally come from the same individual as the engineered cells, but this is impractical in most scenarios. The extent to which inter-individual variation is captured in SCG data and how well this variation is modeled in engineered systems is still being explored. In general, statistical tools that assess reproducibility, correct technical noise, and quantify similarity will improve the ability to assess accuracy and precision from single-cell genomic measurements[37-40].

The precision, efficiency, and accuracy of engineering

The directed differentiation of PSCs to defined cell types in 2D, or the reprogramming of one somatic cell type into another, is often inefficient (Fig. 2a). For example, expressing the single transcription factor Ascl1 results in only ~10% of input fibroblasts becoming induced neuronal cells (iNs). Previous efforts to characterize engineered cells have generally focused on using markers to determine if the target cell has been generated. In contrast, minimal attention has been paid to 'off-target' cell types because it is difficult to predict the appropriate off-target markers. As an alternative, scRNA-seq can provide an unbiased sample of cellular heterogeneity at different time points during reprogramming. For iNs, scRNA-seq analysis showed that myocyte fate emerged during reprogramming and was more abundant than the target neuronal cell fate, though both cell lineages expressed the Tau-Egfp selection marker[41]. When Ascl1 was combined with additional neuron-specific factors, most of the Tau-Egfp positive cells acquired a neuronal fate. Similarly, undesired cell types were detected by scRNA-seq upon iPSC differentiation to cardiomyocytes[42].

These observations indicate that one major reason for reprogramming inefficiency is the lack of precision of the protocol, which provides opportunities to logically engineer the differentiation process. For example, Loh and colleagues mapped the cell fates that emerge during mesoderm development from pluripotency, and identified extrinsic signals that correlate with bifurcating lineage

choices[43]. In this way, they were able to block unwanted lineages and steer the differentiation path towards specific bone and heart progenitors with higher efficiency. Insights from single-cell transcriptomics thus provide strategies to monitor cell fates and to increase efficiency and precision during directed differentiation, by enhancing target fates or supressing off-target fates[44,45] (Fig. 2b).

The targeted fates that emerge from tissue engineering may not exactly represent the cells found in primary human tissue. scRNA-seq can be used to quantify the accuracy of each differentiated cell type. As an example, Le Manno et al. profiled *in vitro* human stem-cell derived dopaminergic neurons, which clustered into 14 molecularly distinct populations, and used machine learning to quantify their similarity to *in vivo* dopaminergic neuron "prototypes" in a human fetal midbrain atlas[46]. The authors found substantial similarity between the *in vitro* and *in vivo* cells, especially in key developmental drivers, but also noted differences in global expression patterns.

Cells in the dish may also take different differentiation paths than cells that develop in the body. Some protocols aim to recapitulate a native developmental sequence by applying a progressive regimen of inductive molecules to differentiating cells. Other directed differentiation protocols bypass certain developmental intermediates to arrive at a mature state. To determine what path cells actually take, scRNA-seq data can be used to computationally reconstruct differentiation trajectories[25,47,48]. Based on the similarity of their single-cell transcriptomes, cells can be placed along a temporal progression, or 'pseudotime' axis. (Pseudotime reflects an ordering of transcriptome states along a process, rather than the actual relative timing of each step.) Since cells in a sampled population undergo differentiation at different rates, pseudotime trajectories can be inferred from a single snapshot, or the approach can be extended to multiple sampling time points in a differentiation experiment. Using this approach, Briggs et al. recently showed that both TF overexpression and growth factor regimens can generate very similar motor neuron states despite the fact that cells progress along different differentiation paths[49]. It will be interesting for the field to pursue the question of whether taking such different routes leads to subtle differences in gene expression or epigenetic memories, or has implications for the use of the engineered cells in disease modeling and therapeutics.

2D cultures lack some of the cell-cell communication that is important for developmental and homeostatic processes, and self-organizing 3D tissues have emerged as powerful models that recover more complex interactions[50]. Organoids contain cell types, morphologies and functions that resemble the corresponding primary tissue, but the accuracy of this recapitulation has been difficult to quantify. Most studies compare organoids to primary tissue using bulk transcriptomics[13,15,51], which lacks cell-type resolution. Using scRNA-seq, we recently uncovered a remarkable correspondence between human organoids and the in vivo fetal differentiation programs of their component cell types[52,53]. Of the genes that vary between progenitors and neurons in the fetal cortex, more than 85% were significantly correlated (>0.4 Pearson's r) with the corresponding differentiation trajectory in cerebral organoids. For liver organoids, hepatocyte-like transcriptome states in the organoid had a maximum of 85% similarity to fetal hepatocytes compared to a maximum of 60% in 2D liver monocultures, which we linked to interlineage signalling in the liver organoid[53]. The analyses also identified gene expression differences between fetal and organoid cells that were related to media components or the absence of particular cell lineages in the organoids.

Detecting these differences can guide strategies to improve the organoid culture protocol as discussed below (reverse engineering). Many differences can arise from comparing organoids and tissues that are at different developmental states. Higher-throughput SCG methods are enabling more sampling of human organoids over time[54]. For example, Quadrato et al. measured the transcriptome of 82,291 individual cells from 31 brain organoids at two time points (3 and 6 months), allowing them to identify diverse cell populations from different brain regions over time, and to report substantial batch-to-batch variation in organoid cell composition. New methods that decrease the cost of sequencing per cell, coupled with strategies to multiplex samples[55], will help measure cells over time-courses and in multiple environmental conditions. This will help disentangle the influence of cell-intrinsic versus environmental effects on the accuracy of the engineered cells, and provide an enhanced understanding of the capacity of these exciting models to recapitulate human physiology.

Reverse engineering specific cell types

Reverse engineering is the process of disassembling an object to understand how it works in order to recreate or enhance it. Single-cell genomics provides new strategies for reverse engineering

tissues by cataloging component parts and predicting how to recreate them. A major goal is to predict the combination of TFs that can generate specific cell types and subtypes. Historically, educated guesses based on expression in bulk samples helped prioritize TFs for small-scale combinatorial screens. The four 'Yamanaka factors' that reprogram somatic cells into PSCs were identified by functionally testing 24 candidates, chosen in part by enriched expression in embryonic stem cells[56]. Similar strategies were used to identify factors to generate neurons from fibroblasts[57]. scRNA-seq can now improve prediction by increasing the resolution of cell states; it can be used to identify the TFs expressed in specific cell subtypes and at branching points along a differentiation path. Computational frameworks[58] can predict TF combinations to differentiate input cells to a target cell subtype.

Pooled screens are more efficient than arrayed screens for testing TF combinations. In the context of pooled cDNA overexpression libraries[59], scRNA-seq can be used to identify the particular overexpressed TFs that lead to emergent cell fates, which can be assessed against a reference cell-type atlas (Fig. 3a). A similar approach can be applied to identify and test signalling molecules that induce differentiation or maturation. This strategy was recently employed for differentiating myoblasts to myocytes; scRNA-seq analysis predicted that modulating insulin and BMP signalling pathways could enhance the MYOD-mediated reprogramming of fibroblast to myocytes. Addition of these signalling molecules enhanced the reprogramming efficiency five-fold[36].

Spatiotemporal triggers for lineage specification are commonly provided by an organized niche microenvironment[60,61]. The description of niches using scRNA-seq can identify signals for maintaining an organ-specific stem cell population or differentiating cells towards a mature cell type[62]. These signalling inputs often have defined spatial locations within a complex tissue. Most widely used scRNA-seq protocols disrupt the spatial integrity of the tissue, however these approaches can be combined with emerging spatial transcriptomic methods (recently reviewed in [63,64]), such as sequential single-molecule fluorescence in situ hybridization[65], to resolve the likely locations of cell states. Recently, Medaglia and colleagues developed the NICHE-seq method, which combines photoactivatable fluorescent reporters, microscopy and scRNA-seq to determine the cellular and molecular composition of immune niches, and can in principle be applied to other niches[66]. The advantage of this protocol is that it is measures the entire transcriptome, rather than a set of a priori determined RNA targets.

In addition to stem cell niches, the functional organization of a tissue is inherently under spatial regulation. Even seemingly homogenous cell populations, such as hepatocytes and enterocytes, exhibit dramatic gene expression differences depending on their location within the tissue's repeating anatomical units (liver lobules[67] or intestinal villi[68]). These cells experience concentration gradients of oxygen, nutrients, morphogens, bacteria and other factors. The cells themselves play a role in shaping the gradients, creating a complex system that is difficult to emulate *in vitro*. Substantial effort is being invested in reconstructing graded microenvironments within various microfluidic or matrix scaffolds, as the microenvironments seem to be critical for generating physiological functions[69,70]. scRNA-seq can guide the structural design of culture units within microfluidic devices or other scaffolds by making it possible to assess spatially-dependent cell states. As spatial transcriptomics methods advance and spread[63,64,71] it will become feasible to explore intercellular interactions during organoid morphogenesis and to learn about potential deficiencies of organoid tissue patterning, niche development and cellular maturation due to missing or misplaced microenvironmental signals (Fig. 3b).

Perturbation, lineage tracing, and epigenomic methods

Single-cell transcriptome measurements are used to dissect cell differentiation trajectories and correlate these with natural development. However, these measurements are descriptive and may only provide hypotheses about underlying mechanisms. In contrast, new methods that couple CRISPR/Cas9 screening with scRNA-seq readouts can systematically assess the effect of genetic perturbations on gene expression in high throughput. Similar to a pooled screen, large numbers of perturbations can be tested at once; however, each cell acts as an independent experiment because it only includes one or a few perturbations, and the readout is the full transcriptome rather than a single selected phenotype. This strategy has been used to investigate regulatory circuits controlling myeloid cell differentiation[72], unfolded protein response[73] and T-cell receptor activation[74].

CRISPR/Cas9 knockout screens will also help elucidate the genes that are necessary for the emergence of specific subtypes and identify the roadblocks that lead to inefficient differentiation and alternative fates (Fig. 3c). At this point, single-cell perturbation methods have only been applied to relatively homogenous cell populations. Transitioning to a heterogeneous tissue culture will require screening fewer genes or pushing the throughput of current scRNA-seq methods in order to sufficiently sample each perturbation in each cell state. Alternatively, an increase in cell throughput might be achieved through sequencing a targeted set of informative transcripts instead of the full transcriptome. Beside loss of function screens, a catalytically inactive version of Cas9 can be combined with transcriptional effectors that either activate or repress transcription of endogenous genes[75]. This approach can be used to identify non-coding regulatory regions that control differentiation towards target cell types[76]. There is immense potential for scRNA-seq-coupled genetic screens to understand mechanisms controlling cell differentiation, barriers to plasticity, and the organization of cells in 3D environments.

Other methods in the single-cell toolbox can help stem cell biologists tackle how cells behave in time. Using scRNA-seq data, computational approaches can order cells in pseudotime and reveal potential lineage bifurcations [25,47,48,77], they can map cell states across time scales[78], and can predict the directionality of lineage progression based on the relative abundance of spliced and unspliced transcripts[79]. However, these methods are based on indirect inference. Complementary methods have been developed to perform highly multiplexed fate mapping and lineage tracing directly in single cells. Two promising classes of these methods either introduce barcoded mRNAs through viral or transposon libraries[80,81] or use Cas9-mediated DNA mutations ('scars') [82,83]. The latter approach using inducible Cas9 systems offers the advantage of creating evolvable barcodes for tracing lineage trees rather than mapping cell fates (Fig. 3d). Further, exciting new methods that allow *in situ* readout of CRISPR-introduced DNA modifications can allow lineage relationships to be discerned while preserving the cells' spatial relationships[84]. Together these studies demonstrate the tremendous opportunities for using lineage-coupled single-cell transcriptomic methods to analyze lineage decisions in human organoids and differentiation events during cell reprogramming.

The majority of SCG studies on engineered cells thus far have focused on the transcriptome, however, it will be important to characterize engineered cells at the level of chromatin organization and epigenetic marks. Towards this aim, methods have been developed to profile accessible chromatin, chromatin looping, and DNA methylation at single-cell resolution[85]. These methods can be used to generate cell atlases similar to what has been accomplished with transcriptomics, as well as to study cell reprogramming. One outstanding question is whether similar cell types or states that arise from two distinct differentiation routes (e.g. direct reprogramming vs. differentiation through all developmental intermediates[49]) coalesce transcriptionally, with equivalent DNA methylation patterns and chromatin organization. This will be particularly interesting for heritable epigenetic marks (e.g. DNA methylation), as disturbances in epigenome maintenance underlie age-related disease[85]. Notably, it was recently suggested that direct differentiation of fibroblasts to neurons preserves aging phenotypes that are erased during reprogramming first to pluripotency and then to neurons[86]. Integrating perturbation, lineage tracing, and epigenomic measures in single engineered cells will connect the genome with its function and enhance computational models of cell fates in health and disease.

Implications for disease modeling and therapies

Induced PSCs (iPSCs) were anticipated to herald an era of personalized medicine by providing patient-specific cell types for disease modeling, drug screening, gene therapy and transplantation[87]. To model disease, specific mutations can also now be introduced into iPSCs using CRISPR/Cas9 genome editing, followed by differentiation into the target cell or tissue type. The high expectations for the field are starting to be met by the development of numerous stem-cell-based cellular disease models and potential therapeutic approaches[88], and 3D organoids are being used to model neurodevelopmental disorders, autism, cystic fibrosis, and metabolic disorders[89]. A mutation (or infection) can have distinct effects on different cells in a heterogeneous 2D cell culture or complex 3D organoid, and disease phenotypes can manifest through effects on a small cell subpopulation. Single-cell transcriptomics provides an unbiased way to search for disease-associated phenotypes in all subpopulations[90]. In addition, integration of genome-wide association data with scRNA-seq can help

to identify the likely cell type impacted by disease-associated genetic variants[91,92]. It is still the very early days of integration between single-cell genomics and *in vitro* disease modeling, but the field is developing rapidly.

One major bottleneck in disease modeling and drug screening is the need for culturing and preparing sequencing libraries from multiple patients and replicates. However, cell populations from different individuals can actually be cultured and sequenced together, and then demultiplexed *in silico* based on RNA sequences that vary uniquely between individuals[93]. This scRNA-seq approach, combined with sample multiplexing[55], might minimize the variation in culture conditions across healthy and disease cell lines (or drug treatment vs. control), and it can also be used to understand cell autonomous and non-cell autonomous effects of disease phenotypes. Ongoing efforts to establish resources for many patient-derived iPSC lines (HipSci, SSCB, Spanish Stem Cell Bank, HSCI iPS Core) will greatly facilitate this undertaking.

There are many cell transplantation therapies now in clinical trials, including iPSC-derived dopaminergic cells to treat Parkinson's disease[8], autologous retinal cells to treat age related macular degeneration[10], and immunotherapy to treat cancer[94,95], where single-cell genomics could have a direct impact. scRNA-seq can be applied to assess the heterogeneity, purity, and potential safety of each batch of these therapeutics. Initial analyses of heterogeneity within and between batches of engineered cells using scRNA-seq could be used to identify quality control biomarkers and inform lower cost methods (e.g. immunohistochemistry) to survey potential off-targets. Furthermore, failures in cell therapy clinical trials can lead to tragic outcomes[96] and SCG methods could be used to monitor heterogeneity and understand what goes right and wrong during future trials of promising cell transplantation therapies.

Single-cell DNA and RNA sequencing has already been used to explore the clonal heterogeneity and evolution of cancer tissue[97,98]. Advances in cancer organoid culture will enable the study of tumour heterogeneity *in vitro* with these technologies, facilitating a move past the "one size fits all" treatment approach. One question is which cell types within a complex tumor should be targeted in order to achieve maximally favorable treatment outcomes. The cancer stem cell (CSC) concept—that some tumor cells can be hierarchically organized similar to somatic stem cells—sparked great excitement in the field of cancer biology[99]. The promise of the CSC concept was to eradicate the "beating heart" of the tumor by targeting the tumor at its source[100]. Extensive research efforts in the last years have demonstrated, however, that cancer cells are plastic and even non-CSCs can acquire CSC properties[101,102]. SCG methods have the potential to unravel how cancer cells traverse cell states during processes such as cancer cell dissemination, epithelial-to-mesenchymal transition and acquisition of drug resistance, and patient-specific tumor organoids will enable the customized study and targeting of tumor cells[103,104].

Outlook

The single-cell genomic and tissue engineering fields are in a phase of rapid growth. We expect that the coming years will bring integrated technologies that combine high-throughput, lineage-coupled, and spatially resolved single-cell multi-omic measurements. These new methods will help to generate more accurate, precise and efficient cultures of engineered cells that can be used for transplantation. Applied to patient-derived organoid models of disease and cancer, this merger offers exciting opportunities to generate organ-level computational models of healthy human development, to understand what goes wrong in disease, and to predict treatment outcomes.

Acknowledgements

We thank Keisuke Sekine and Tal Nawy for reading the manuscript and for their very thoughtful guidance. Funding for this work was provided by the Max Planck Society (GC, DW, BT) and a European Research Council Starting Grant (ID:758877, BT).

Competing Financial Interests Statement

The authors have no competing financial interests.

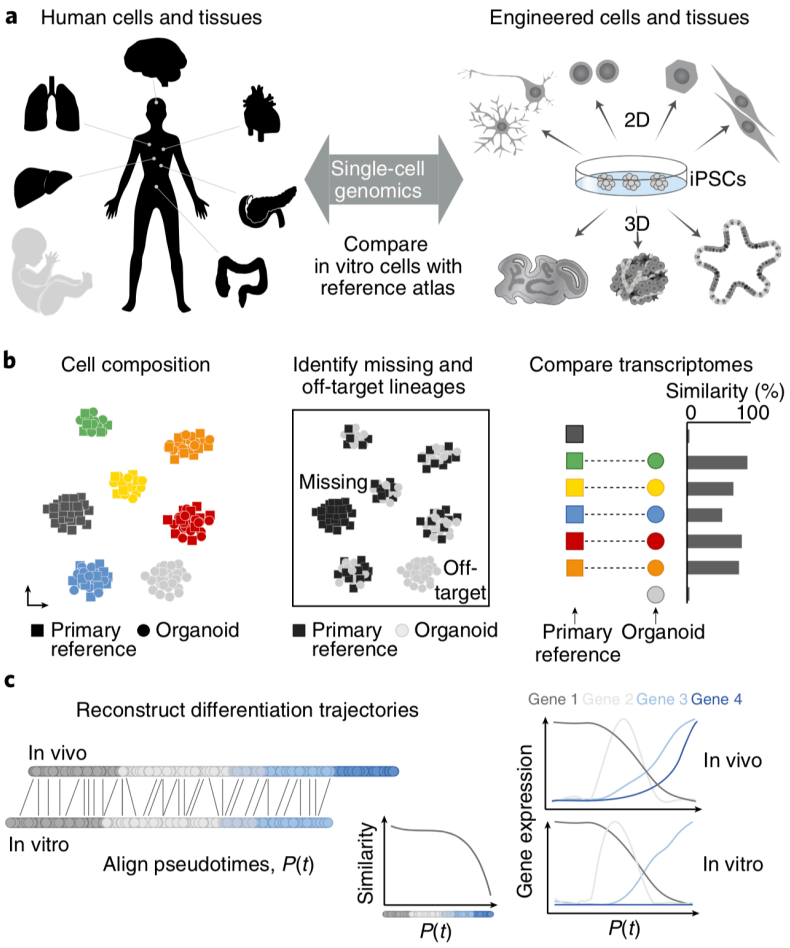

Figure 1: A human cell atlas is an optimum reference for cell and tissue engineering. a) Single-cell genomic (SCG) methods are being used to make comprehensive atlases of cell type diversity for diverse human organ systems. In parallel, SCG approaches can be applied to dissect heterogeneity arising during cell and tissue engineering. b,c, Primary and engineered cells can then be compared in terms of (b) cell composition and transcriptome similarities or (c) across differentiation pathways and over time.

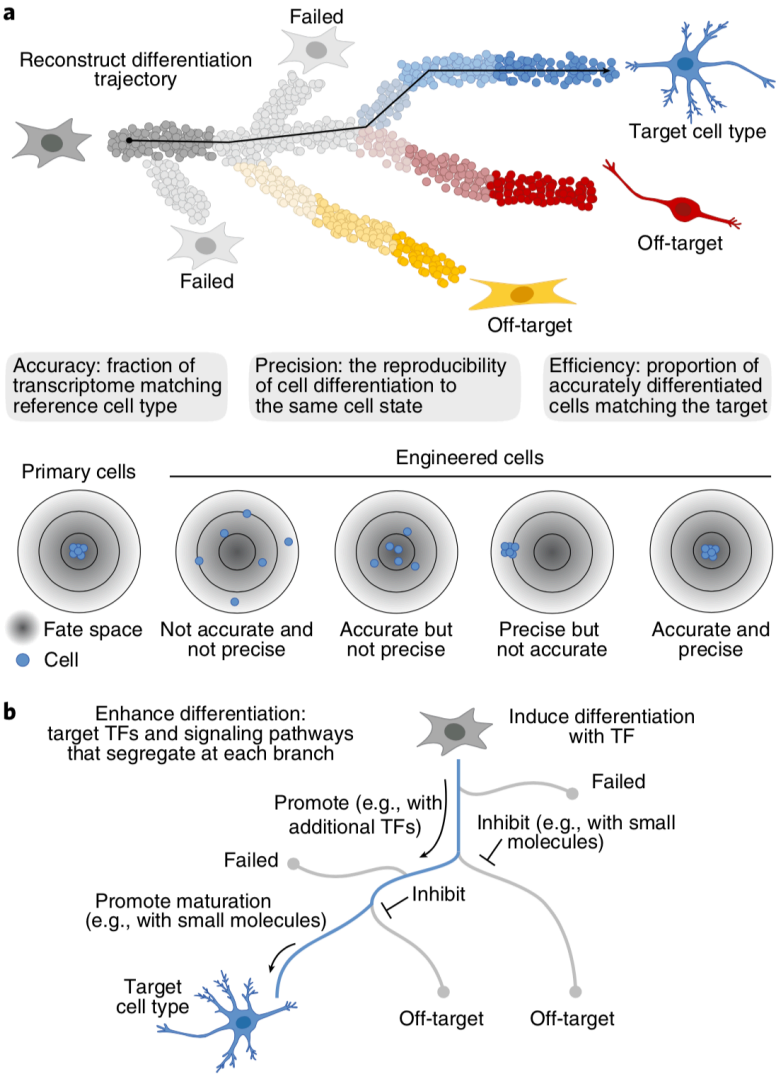

Figure 2: Single-cell genomics to assess and enhance accuracy, precision, and efficiency during directed differentiation a) Single-cell transcriptomics can be used to reconstruct differentiation paths and illuminate failed differentiation events or off-target cell fates that can emerge alongside the target cell type during directed differentiation. The final mixture of cells in a differentiation experiment (represented as endpoints in the reconstructed pseudotemporal trajectory) can be compared to a reference atlas, and the accuracy, precision, and efficiency of the engineering process can be quantified. One can imagine the fate space as a dartboard, and an accurate, precise, and efficient engineering protocol will generate cells tightly clustered around the target fate in the center. b) Single-cell transcriptome data can be used to identify the transcription factors and signaling pathways that segregate at each branch, and this information can be used to enhance differentiation of the target cell type.

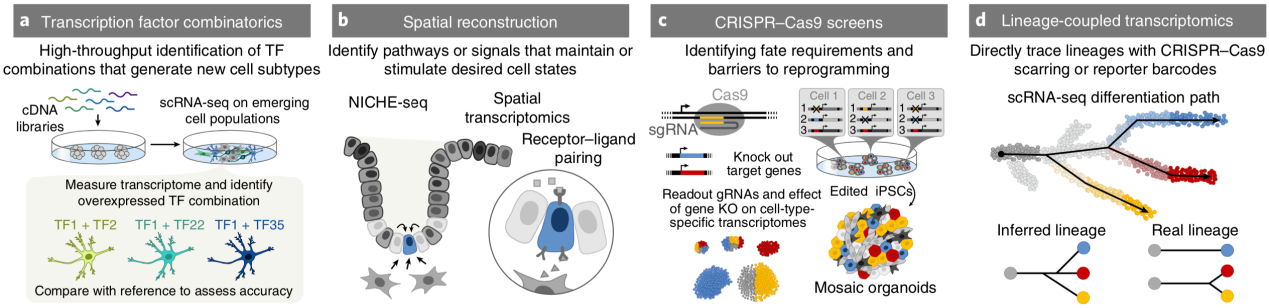

Figure 3: Emerging methodologies that will guide cell and tissue engineering. a, Identification of effective TF combinations. b, Spatial reconstruction of differentiation pathways. c, CRISPR–Cas9 screening for genes central to cell fate determination. d, Lineage tracing by transcriptomics.

REFERENCES

1 Picelli, S. Single-cell RNA-sequencing: The future of genome biology is now. *RNA Biol*, 1-14, doi:10.1080/15476286.2016.1201618 (2016).

2 Tanay, A. & Regev, A. Scaling single-cell genomics from phenomenology to mechanism. *Nature* **541**, 331-338, doi:10.1038/nature21350 (2017).

3 Pang, Z. P. *et al.* Induction of human neuronal cells by defined transcription factors. *Nature* **476**, 220-223, doi:10.1038/nature10202 (2011).

4 Takahashi, K. & Yamanaka, S. A decade of transcription factor-mediated reprogramming to pluripotency. *Nat Rev Mol Cell Biol* **17**, 183-193, doi:10.1038/nrm.2016.8 (2016).

5 Karow, M. *et al.* Reprogramming of pericyte-derived cells of the adult human brain into induced neuronal cells. *Cell stem cell* **11**, 471-476, doi:10.1016/j.stem.2012.07.007 (2012).

6 Graf, T. Historical origins of transdifferentiation and reprogramming. *Cell stem cell* **9**, 504-516, doi:10.1016/j.stem.2011.11.012 (2011).

7 Xu, J., Du, Y. & Deng, H. Direct lineage reprogramming: strategies, mechanisms, and applications. *Cell stem cell* **16**, 119-134, doi:10.1016/j.stem.2015.01.013 (2015).

8 Kriks, S. *et al.* Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature* **480**, 547-551, doi:10.1038/nature10648 (2011).

9 Chong, J. J. *et al.* Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. *Nature* **510**, 273-277, doi:10.1038/nature13233 (2014).

10 Mandai, M. *et al.* Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration. *The New England journal of medicine* **376**, 1038-1046, doi:10.1056/NEJMoa1608368 (2017).

11 Eiraku, M. *et al.* Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472**, 51-56, doi:10.1038/nature09941 (2011).

12 Lancaster, M. A. *et al.* Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373-379, doi:10.1038/nature12517 (2013).

13 Takebe, T. *et al.* Generation of a vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature protocols* **9**, 396-409, doi:10.1038/nprot.2014.020 (2014).
14 McCracken, K. W. *et al.* Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* **516**, 400-404, doi:10.1038/nature13863 (2014).
15 Spence, J. R. *et al.* Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* **470**, 105-109, doi:10.1038/nature09691 (2011).
16 Takasato, M. *et al.* Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature* **526**, 564-568, doi:10.1038/nature15695 (2015).
17 Birey, F. *et al.* Assembly of functionally integrated human forebrain spheroids. *Nature* **545**, 54-59, doi:10.1038/nature22330 (2017).
18 Bagley, J. A., Reumann, D., Bian, S., Levi-Strauss, J. & Knoblich, J. A. Fused cerebral organoids model interactions between brain regions. *Nature methods* **14**, 743-751, doi:10.1038/nmeth.4304 (2017).
19 Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nature methods* **6**, 377-382, doi:10.1038/nmeth.1315 (2009).
20 Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202-1214, doi:10.1016/j.cell.2015.05.002 (2015).
21 Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187-1201, doi:10.1016/j.cell.2015.04.044 (2015).
22 Habib, N. *et al.* Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* **353**, 925-928, doi:10.1126/science.aad7038 (2016).
23 Shekhar, K. *et al.* Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell* **166**, 1308-1323 e1330, doi:10.1016/j.cell.2016.07.054 (2016).
24 Rosenberg, A. B. *et al.* Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*, doi:10.1126/science.aam8999 (2018).
25 Bendall, S. C. *et al.* Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell* **157**, 714-725, doi:10.1016/j.cell.2014.04.005 (2014).
26 Villani, A. C. *et al.* Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, doi:10.1126/science.aah4573 (2017).
27 Pollen, A. A. *et al.* Molecular Identity of Human Outer Radial Glia during Cortical Development. *Cell* **163**, 55-67, doi:10.1016/j.cell.2015.09.004 (2015).
28 Nowakowski, T. J. *et al.* Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318-1323, doi:10.1126/science.aap8809 (2017).
29 Rozenblatt-Rosen, O., Stubbington, M. J. T., Regev, A. & Teichmann, S. A. The Human Cell Atlas: from vision to reality. *Nature* **550**, 451-453, doi:10.1038/550451a (2017).
30 Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature biotechnology* **36**, 411-420, doi:10.1038/nbt.4096 (2018).

31 Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nature biotechnology* **36**, 421-427, doi:10.1038/nbt.4091 (2018).
32 Butler, A. & Satija, R. Integrated analysis of single cell transcriptomic data across conditions, technologies, and species. *bioRxiv*, doi:10.1101/164889 (2017).
33 Kiselev, V. Y., Yiu, A. & Hemberg, M. scmap - A tool for unsupervised projection of single cell RNA-seq data. *bioRxiv*, doi:10.1101/150292 (2017).
34 van den Brink, S. C. *et al.* Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nature methods* **14**, 935-936, doi:10.1038/nmeth.4437 (2017).
35 Alpert, A., Moore, L. S., Dubovik, T. & Shen-Orr, S. S. Alignment of single-cell trajectories to compare cellular expression dynamics. *Nature methods*, doi:10.1038/nmeth.4628 (2018).
36 Cacchiarelli, D. *et al.* Aligning single-cell developmental and reprogramming trajectories identifies molecular determinants of reprogramming outcome. *bioRxiv*, doi:10.1101/122531 (2017).
37 Stegle, O., Teichmann, S. A. & Marioni, J. C. Computational and analytical challenges in single-cell transcriptomics. *Nature reviews. Genetics* **16**, 133-145, doi:10.1038/nrg3833 (2015).
38 Cahan, P. *et al.* CellNet: network biology applied to stem cell engineering. *Cell* **158**, 903-915, doi:10.1016/j.cell.2014.07.020 (2014).
39 Symmons, O. & Raj, A. What's Luck Got to Do with It: Single Cells, Multiple Fates, and Biological Nondeterminism. *Molecular cell* **62**, 788-802, doi:10.1016/j.molcel.2016.05.023 (2016).
40 Soneson, C. & Robinson, M. D. Bias, robustness and scalability in single-cell differential expression analysis. *Nature methods*, doi:10.1038/nmeth.4612 (2018).
41 Treutlein, B. *et al.* Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* **534**, 391-395, doi:10.1038/nature18323 (2016).
42 Bargaje, R. *et al.* Cell population structure prior to bifurcation predicts efficiency of directed differentiation in human induced pluripotent cells. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 2271-2276, doi:10.1073/pnas.1621412114 (2017).
43 Loh, K. M. *et al.* Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell stem cell* **14**, 237-252, doi:10.1016/j.stem.2013.12.007 (2014).
44 Kee, N. *et al.* Single-Cell Analysis Reveals a Close Relationship between Differentiating Dopamine and Subthalamic Nucleus Neuronal Lineages. *Cell stem cell* **20**, 29-40, doi:10.1016/j.stem.2016.10.003 (2017).
45 Yao, Z. *et al.* A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. *Cell stem cell* **20**, 120-134, doi:10.1016/j.stem.2016.09.011 (2017).
46 La Manno, G. *et al.* Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells. *Cell* **167**, 566-580 e519, doi:10.1016/j.cell.2016.09.027 (2016).
47 Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature biotechnology* **32**, 381-386, doi:10.1038/nbt.2859 (2014).

48 Haghverdi, L., Buttner, M., Wolf, F. A., Buettner, F. & Theis, F. J. Diffusion pseudotime robustly reconstructs lineage branching. *Nature methods* **13**, 845-848, doi:10.1038/nmeth.3971 (2016).
49 Briggs, J. A. *et al.* Mouse embryonic stem cells can differentiate via multiple paths to the same state. *eLife* **6**, doi:10.7554/eLife.26945 (2017).
50 Clevers, H. Modeling Development and Disease with Organoids. *Cell* **165**, 1586-1597, doi:10.1016/j.cell.2016.05.082 (2016).
51 Qian, X. *et al.* Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell* **165**, 1238-1254, doi:10.1016/j.cell.2016.04.032 (2016).
52 Camp, J. G. *et al.* Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 15672-15677, doi:10.1073/pnas.1520760112 (2015).
53 Camp, J. G. *et al.* Multilineage communication regulates human liver bud development from pluripotency. *Nature* **546**, 533-538, doi:10.1038/nature22796 (2017).
54 Quadrato, G. *et al.* Cell diversity and network dynamics in photosensitive human brain organoids. *Nature* **545**, 48-53, doi:10.1038/nature22047 (2017).
55 Stoeckius, M. *et al.* Cell "hashing" with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *bioRxiv*, doi:10.1101/237693 (2017).
56 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).
57 Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035-1041, doi:nature08797 [pii]
10.1038/nature08797 [doi] (2010).
58 Rackham, O. J. *et al.* A predictive computational framework for direct reprogramming between human cell types. *Nature genetics* **48**, 331-335, doi:10.1038/ng.3487 (2016).
59 Tsunemoto, R. *et al.* Diverse reprogramming codes for neuronal identity. *Nature*, doi:10.1038/s41586-018-0103-5 (2018).
60 Blanpain, C., Lowry, W. E., Geoghegan, A., Polak, L. & Fuchs, E. Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell* **118**, 635-648, doi:10.1016/j.cell.2004.08.012 (2004).
61 Morrison, S. J. & Spradling, A. C. Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell* **132**, 598-611, doi:10.1016/j.cell.2008.01.038 (2008).
62 Li, L. *et al.* Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell stem cell* **20**, 858-873 e854, doi:10.1016/j.stem.2017.03.007 (2017).
63 Lein, E., Borm, L. E. & Linnarsson, S. The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science* **358**, 64-69, doi:10.1126/science.aan6827 (2017).
64 Crosetto, N., Bienko, M. & van Oudenaarden, A. Spatially resolved transcriptomics and beyond. *Nature reviews. Genetics* **16**, 57-66, doi:10.1038/nrg3832 (2015).

65 Shah, S., Lubeck, E., Zhou, W. & Cai, L. In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. *Neuron* **92**, 342-357, doi:10.1016/j.neuron.2016.10.001 (2016).
66 Medaglia, C. *et al.* Spatial reconstruction of immune niches by combining photoactivatable reporters and scRNA-seq. *Science* **358**, 1622-1626, doi:10.1126/science.aao4277 (2017).
67 Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352-356, doi:10.1038/nature21065 (2017).
68 Moor, A. E. *et al.* Spatial reconstruction of single enterocytes uncovers broad zonation along the intestinal villus axis. *bioRxiv*, doi:10.1101/261529 (2018).
69 Bhatia, S. N. & Ingber, D. E. Microfluidic organs-on-chips. *Nature biotechnology* **32**, 760-772, doi:10.1038/nbt.2989 (2014).
70 Gjorevski, N. *et al.* Designer matrices for intestinal stem cell and organoid culture. *Nature* **539**, 560-564, doi:10.1038/nature20168 (2016).
71 Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**, doi:10.1126/science.aat5691 (2018).
72 Jaitin, D. A. *et al.* Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell* **167**, 1883-1896 e1815, doi:10.1016/j.cell.2016.11.039 (2016).
73 Adamson, B. *et al.* A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* **167**, 1867-1882 e1821, doi:10.1016/j.cell.2016.11.048 (2016).
74 Datlinger, P. *et al.* Pooled CRISPR screening with single-cell transcriptome readout. *Nature methods* **14**, 297-301, doi:10.1038/nmeth.4177 (2017).
75 Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442-451, doi:10.1016/j.cell.2013.06.044 (2013).
76 Montalbano, A., Canver, M. C. & Sanjana, N. E. High-Throughput Approaches to Pinpoint Function within the Noncoding Genome. *Molecular cell* **68**, 44-59, doi:10.1016/j.molcel.2017.09.017 (2017).
77 Haghverdi, L., Buettner, F. & Theis, F. J. Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics* **31**, 2989-2998, doi:10.1093/bioinformatics/btv325 (2015).
78 Schiebinger, G. *et al.* Reconstruction of developmental landscapes by optimal-transport analysis of single-cell gene expression sheds light on cellular reprogramming. *bioRxiv*, doi:10.1101/191056 (2017).
79 La Manno, G. *et al.* RNA velocity in single cells. *bioRxiv*, doi:10.1101/206052 (2017).
80 Biddy, B. A., Waye, S. E., Sun, T. & Morris, S. A. Single-cell analysis of clonal dynamics in direct lineage reprogramming: a combinatorial indexing method for lineage tracing. *bioRxiv*, doi:10.1101/127860 (2017).
81 Wagner, D. E. *et al.* Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981-987, doi:10.1126/science.aar4362 (2018).
82 Spanjaard, B. *et al.* Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nature biotechnology* **36**, 469-473, doi:10.1038/nbt.4124 (2018).

83 Farrell, J. A. *et al.* Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, doi:10.1126/science.aar3131 (2018).
84 Frieda, K. L. *et al.* Synthetic recording and in situ readout of lineage information in single cells. *Nature* **541**, 107-111, doi:10.1038/nature20777 (2017).
85 Smallwood, S. A. *et al.* Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nature methods* **11**, 817-820, doi:10.1038/nmeth.3035 (2014).
86 Mertens, J. *et al.* Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects. *Cell stem cell* **17**, 705-718, doi:10.1016/j.stem.2015.09.001 (2015).
87 Dekkers, J. F. *et al.* Characterizing responses to CFTR-modulating drugs using rectal organoids derived from subjects with cystic fibrosis. *Sci Transl Med* **8**, 344ra384, doi:10.1126/scitranslmed.aad8278 (2016).
88 Trounson, A. & McDonald, C. Stem Cell Therapies in Clinical Trials: Progress and Challenges. *Cell stem cell* **17**, 11-22, doi:10.1016/j.stem.2015.06.007 (2015).
89 Dutta, D., Heo, I. & Clevers, H. Disease Modeling in Stem Cell-Derived 3D Organoid Systems. *Trends in molecular medicine* **23**, 393-410, doi:10.1016/j.molmed.2017.02.007 (2017).
90 Keren-Shaul, H. *et al.* A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell* **169**, 1276-1290 e1217, doi:10.1016/j.cell.2017.05.018 (2017).
91 Park, J. *et al.* Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science*, doi:10.1126/science.aar2131 (2018).
92 Calderon, D. *et al.* Inferring Relevant Cell Types for Complex Traits by Using Single-Cell Gene Expression. *American journal of human genetics* **101**, 686-699, doi:10.1016/j.ajhg.2017.09.009 (2017).
93 Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nature biotechnology*, doi:10.1038/nbt.4042 (2017).
94 Kalos, M. *et al.* T cells with chimeric antigen receptors have potent antitumor effects and can establish memory in patients with advanced leukemia. *Sci Transl Med* **3**, 95ra73, doi:10.1126/scitranslmed.3002842 (2011).
95 Maude, S. L. *et al.* Chimeric antigen receptor T cells for sustained remissions in leukemia. *The New England journal of medicine* **371**, 1507-1517, doi:10.1056/NEJMoa1407222 (2014).
96 Couzin-Frankel, J. Worries, confusion after cancer trial deaths. *Science* **354**, 1211, doi:10.1126/science.354.6317.1211 (2016).
97 Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189-196, doi:10.1126/science.aad0501 (2016).
98 Wang, Y. *et al.* Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155-160, doi:10.1038/nature13600 (2014).
99 Visvader, J. E. & Lindeman, G. J. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nature reviews. Cancer* **8**, 755-768, doi:10.1038/nrc2499 (2008).
100 Clevers, H. The cancer stem cell: premises, promises and challenges. *Nat Med* **17**, 313-319, doi:10.1038/nm.2304 (2011).
101 Gupta, P. B. *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633-644, doi:10.1016/j.cell.2011.07.026 (2011).

102 Batlle, E. & Clevers, H. Cancer stem cells revisited. *Nat Med* **23**, 1124-1134, doi:10.1038/nm.4409 (2017).
103 Boj, S. F. *et al.* Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**, 324-338, doi:10.1016/j.cell.2014.12.021 (2015).
104 Huang, L. *et al.* Ductal pancreatic cancer modeling and drug screening using human pluripotent stem cell- and patient-derived tumor organoids. *Nat Med* **21**, 1364-1371, doi:10.1038/nm.3973 (2015).