nature methods

REVIEW ARTICLE

https://doi.org/10.1038/s41592-018-0113-0

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

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

To understand human development and disease, as well as to regenerate damaged tissues, scientists are working to engineer certain cell types in vitro and to create 3D microenvironments in which cells behave physiologically. Single-cell genomics (SCG) technologies are being applied to primary human organs and to 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 scientists can use single-cell sequencing to optimize human cell and tissue engineering by measuring precision, detecting inefficiencies, and assessing accuracy. We also provide a perspective on how emerging SCG methods can be used to reverse-engineer human cells and tissues and unravel disease mechanisms.

SCG has recently emerged as a field, thanks to the development of an exciting set of tools that enable scientists to explore cellular complexity in cell cultures and tissues and to reconstruct differentiation processes[1,2]. In parallel, enormous progress has been made in the engineering of human cell types and tissues from pluripotent stem cells (PSCs) in culture. In this review, we argue that SCG approaches offer new opportunities to test existing differentiation protocols and their limitations. The resulting inferences could guide strategies to reverse-engineer human organs and explore disease mechanisms.

2D and 3D strategies to engineer human cell types

PSCs grown in 2D culture can be induced to become specific cell types through a regimen of developmental signaling 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 via 2D approaches have been used to investigate mechanisms that control differentiation and barriers to plasticity; in some cases they have been used to model disease, and even to regenerate damaged tissues[8–10].

Of course, human tissues are 3D. They are composed of many different cell types that signal to each other and coordinate functions at the tissue and organ levels. 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 eye[11], brain[12], liver[13], stomach[14], intestine[15], kidney[16], and others. Organoid technologies offer great promise for the modeling of human development and disease, and scientists are starting to assemble modular developing units of tissue (such as ventral and dorsal telencephalon[17,18]) to create controlled intertissue interactions. The ultimate goal is to connect multiple organs into systems for the study of human physiology.

However, there are many bottlenecks and challenges in human cell and tissue engineering. It is often unclear how similar an engineered cell or tissue is to its 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 SCG can address these shortcomings.

Single-cell atlases are optimal references for engineering

It is incredible that the first demonstration of transcriptome sequencing from a single cell was published only 9 years ago[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 catalog 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 for diagnosing, monitoring, and treating disease. Atlas efforts can also serve as a reference for cell and tissue engineering, and enable 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 been useful for benchmarking 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 with those of the primary counterparts from the reference atlas (Fig. 1b). One can quantitate the accuracy (the fraction of the in vitro transcriptome that resembles that of the corresponding primary cell), precision (the ratio of target to off-target lineages in engineered cell cultures), and efficiency (the 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 with the

[1]Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. [2]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany. [3]Department of Biosciences, Technical University Munich, Freising, Germany. *e-mail: gray_camp@eva.mpg.de; Barbara_treutlein@eva.mpg.de

same chemistry and technical platform. However, diverse methods exist for the generation of single-cell transcriptomes, and this can lead to technical differences that confound quantitative comparisons. Computational approaches are being developed that allow datasets generated via different methods or from different species to be integrated[30–32], and that can project clusters from one dataset onto another[33]. Even between experiments conducted 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 and mechanical stress resulting from tissue perfusion or trituration. Optimization of tissue dissociation is therefore a critical aspect of any SCG 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. Alignment of differentiation trajectories can make these comparisons possible[35,36] (Fig. 1c). Finally, an optimal 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 are 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 SCG measurements[37–40].

The precision, efficiency, and accuracy of engineering

The directed differentiation of PSCs to defined cell types in two dimensions, or the reprogramming of one somatic cell type into another, is often inefficient (Fig. 2a). For example, expression of the single transcription factor Ascl1 results in only ~10% of input mouse fibroblasts becoming induced neuronal cells. Previous efforts to characterize engineered cells have generally focused on the use of markers to determine whether 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 induced neuronal cells derived from mouse embryonic fibroblasts, scRNA-seq analysis showed that a 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$^+$ cells acquired a neuronal fate. Similarly, undesired cell types were detected by scRNA-seq after differentiation of induced pluripotent stem cell (iPSCs) 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 et al. 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 toward specific bone and heart progenitors with greater 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 suppressing off-target fates[44,45] (Fig. 2b).

The targeted fates that emerge from cell or tissue engineering might 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, La 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

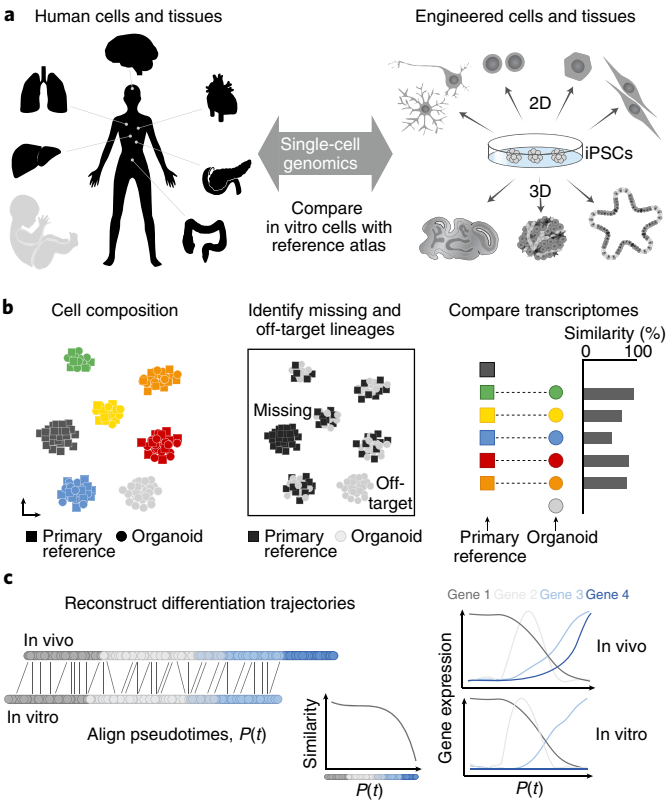

Fig. 1 | A human cell atlas is an optimal reference for cell and tissue engineering. a, SCG methods are being used to make comprehensive atlases of the cell-type diversity of human organ systems. In parallel, SCG approaches can be applied to dissect heterogeneity that arises during cell and tissue engineering.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

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 a 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 cell state. To determine what path cells actually take, one can use scRNA-seq data to computationally reconstruct differentiation trajectories[25,47,48]. On the basis of 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 timing of each step.) Because 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 in the two conditions[49]. It will be interesting for the field to pursue the question of whether such different routes lead to important differences in gene expression or epigenetic memories, or have 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 those in the corresponding primary tissue, but the accuracy of this recapitulation has been difficult to quantify. Most studies compare organoids with primary tissue by means of bulk transcriptomics[13,15,51], which lacks cell-type resolution. Using scRNA-seq, we recently uncovered a remarkable correspondence between the differentiation programs of human organoids and their in vivo fetal counterparts[52,53]. Of the genes that varied between progenitors and neurons in the fetal cortex, more than 85% were significantly correlated (Pearson's $r > 0.4$) 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 with a maximum of 60% in 2D liver monocultures, which we linked to interlineage signaling in the liver organoid[53]. The analyses also identified differences in gene expression between fetal and organoid cells that were related to media components or the absence of particular cell lineages in the organoids.

Detection of these differences can guide strategies to improve the organoid culture protocol, as discussed below (reverse engineering). Many differences can arise between organoids and tissues at different developmental states. Higher-throughput SCG methods are enabling more sampling of human organoids over time[54]. For example, Quadrato et al.[54] measured the transcriptome of 82,291 individual cells from 31 brain organoids at two time points (3 and 6 months after the start of in vitro organoid development), which allowed them to identify diverse cell populations from different brain regions over time and 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 scientists measure cells over different 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 allow for an enhanced understanding of the capacity of these exciting models to recapitulate human physiology.

Reverse engineering of specific cell types

Reverse engineering is the process of disassembling an object to understand how it works in order to recreate or improve it. SCG allows new strategies for the reverse engineering of tissues by enabling researchers to catalog component parts and predict how to recreate them. A major goal is to predict combinations of TFs that can generate specific cell types and subtypes. Historically, educated guesses based on expression in bulk samples have helped researchers prioritize TFs for small-scale combinatorial screens. The four 'Yamanaka factors' that reprogram somatic cells into PSCs were identified through functional testing of 24 candidates, chosen in part on the basis of their 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 such predictions 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 can be 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 by comparison to a reference cell-type atlas (Fig. 3a). A similar approach can be applied to identify and test signaling molecules that induce differentiation or maturation. This strategy was recently used to differentiate myoblasts to myocytes; scRNA-seq analysis predicted that modulation of insulin and BMP (bone morphogenetic protein) signaling pathways could enhance the MYOD-mediated reprogramming of fibroblasts to myocytes. The addition of these signaling molecules improved the reprogramming efficiency fivefold[36].

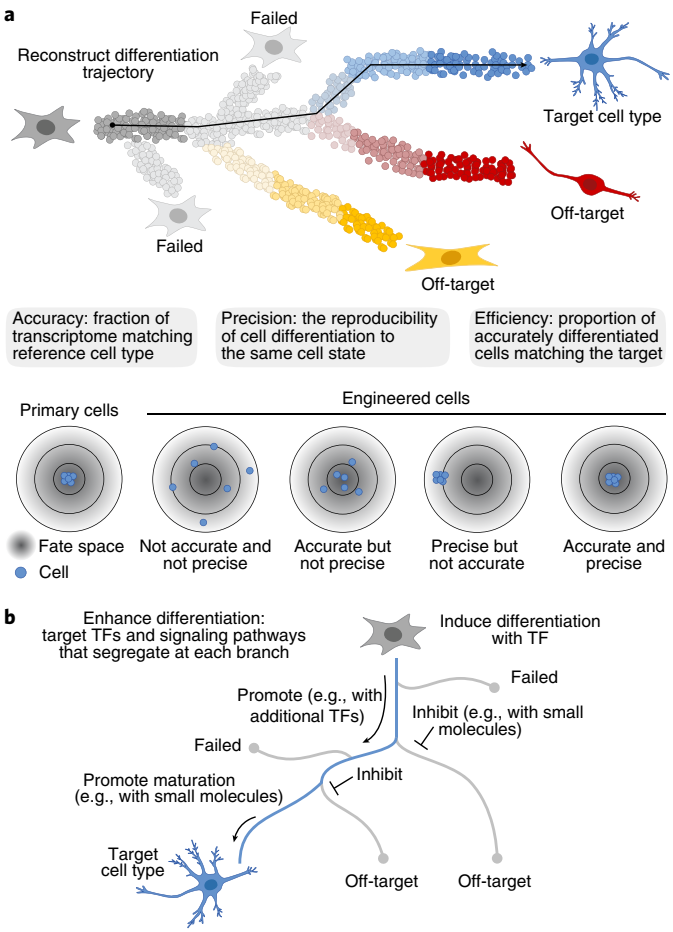

Fig. 2 | Single-cell genomics can be used to assess and enhance accuracy, precision, and efficiency during directed differentiation. a, Single-cell transcriptomics can be used to reconstruct differentiation paths and illuminate failed differentiation events or off-target cell fates that can emerge alongside the target cell type during directed differentiation. The final mixture of cells in a differentiation experiment (represented as endpoints in the reconstructed pseudotemporal trajectory (top)) can be compared to a reference atlas, and the accuracy, precision, and efficiency of the engineering process can be quantified. The images at the bottom depict the fate space imagined as a dartboard; 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 improve differentiation to the target cell type.

Spatiotemporal triggers for lineage specification are commonly provided by an organized niche microenvironment[60,61]. The description of niches by scRNA-seq can identify signals for maintaining an organ-specific stem cell population or differentiating cells toward a mature cell type[62]. These signaling inputs often have defined spatial locations in 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 refs. [63,64]), such as sequential single-molecule fluorescence in situ hybridization[65], to resolve the likely locations of cell states. Recently, Medaglia et al. developed the

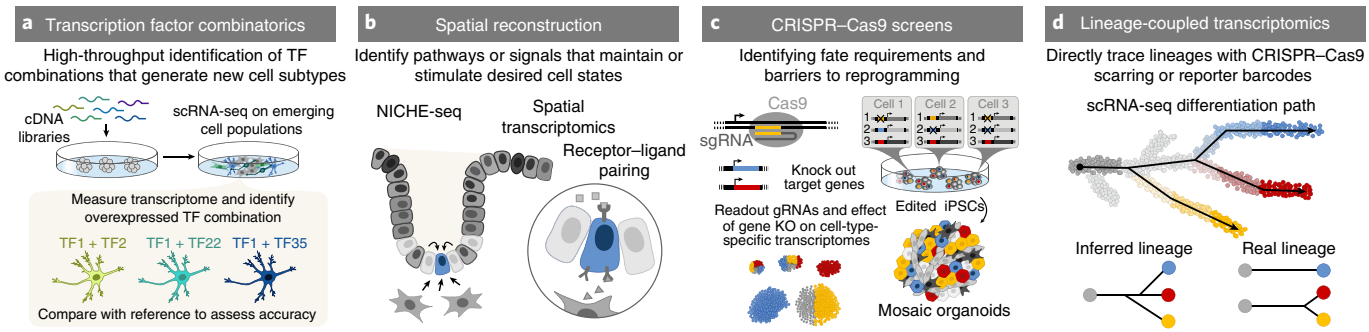

Fig. 3 | Emerging methodologies that will guide cell and tissue engineering. a, Identification of effective TF combinations. **b**, Spatial reconstruction of differentiation pathways. **c**, CRISPR–Cas9 screening for genes central to cell fate determination. **d**, Lineage tracing by transcriptomics.

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 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 homogeneous cell populations, such as hepatocytes and enterocytes, exhibit dramatic differences in gene expression depending on their location within the tissue's repeating anatomical units (liver lobules[67] and intestinal villi[68], respectively). These cells experience concentration gradients of oxygen, nutrients, morphogens, bacteria, and other factors. The cells themselves play a role in shaping the gradients, thereby creating a complex system that is difficult to emulate in vitro. Substantial effort is being invested in reconstructing graded microenvironments in various microfluidic or matrix scaffolds, as the microenvironments seem to be critical for the generation of physiological functions[69,70]. scRNA-seq can guide the structural design of culture units in 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 them with natural development. However, these measurements are descriptive and might 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 at 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 includes only 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 that control 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 been applied only to relatively homogeneous cell populations. The transition to heterogeneous tissue cultures will require screening of fewer genes or an increase in the throughput of current scRNA-seq methods in order for each perturbation in each cell state to be sufficiently sampled. Alternatively, an increase in cell throughput might be achieved through sequencing of a targeted set of informative transcripts instead of the full transcriptome. Besides 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 noncoding regulatory regions that control differentiation toward target cell types[76]. There is immense potential for scRNA-seq-coupled genetic screens to elucidate 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 cell behavior over time. Using scRNA-seq data, computational approaches can order cells in pseudotime and reveal potential lineage bifurcations[25,47,48,77], map cell states across time scales[78], and predict the directionality of lineage progression on the basis of the relative abundance of spliced and unspliced transcripts[79]. However, these methods are based on indirect inference. Complementary methods have been developed for 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 the use of 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, but it also will be important to characterize engineered cells at the level of chromatin organization and epigenetic marks. Toward this end, 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 (for example, direct reprogramming versus 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]. Integration of perturbation, lineage tracing, and epigenomic measures in single engineered cells will connect the genome with its function and improve computational models of cell fates in health and disease.

Implications for disease modeling and therapies

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]. For disease modeling, specific mutations can also now be introduced into iPSCs via CRISPR–Cas9 genome editing, and the iPSCs can then be differentiated 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 could identify the likely cell type(s) affected by disease-associated genetic variants[91,92]. These are still the very early days of integration between SCG and in vitro disease modeling, but the field is developing rapidly.

One major bottleneck in disease modeling and drug screening is the need to culture and prepare sequencing libraries from multiple patients and replicates. However, cell populations from different individuals can actually be cultured and sequenced together and then be demultiplexed in silico on the basis of RNA sequences that are unique to 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 versus control conditions), 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 (e.g., HipSci, HSCI iPS Core,WiCell, NYSCF) will greatly facilitate this undertaking.

There are many cell-transplantation therapies now in clinical trials—for example, the use of 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]—in which SCG 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 by 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 have already been used to explore the clonal heterogeneity and evolution of cancer tissue[97,98]. Advances in cancer organoid culture will allow the study of tumor heterogeneity in vitro with these technologies, thereby facilitating a move past the 'one size fits all' treatment approach. One question is which cell types within a complex tumor should be targeted to achieve maximally favorable treatment outcomes. The cancer stem cell (CSC) concept—the idea 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 a tumor by targeting the tumor at its source[100]. Extensive research efforts in recent 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 allow the customized study and targeting of tumor cells[103,104].

Outlook

The fields of SCG and tissue engineering 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 scientists generate more accurate, precise, and efficient cultures of engineered cells that can be used for transplantation. If applied to patient-derived organoid models of disease and cancer, this merger will offer exciting opportunities to generate organ-level computational models of healthy human development, determine what goes wrong in disease, and predict treatment outcomes.

Received: 10 January 2018; Accepted: 2 August 2018;
Published online: 31 August 2018

References

1. Picelli, S. Single-cell RNA-sequencing: the future of genome biology is now. *RNA Biol.* **14**, 637–650 (2017).
2. Tanay, A. & Regev, A. Scaling single-cell genomics from phenomenology to mechanism. *Nature* **541**, 331–338 (2017).
3. Pang, Z. P. et al. Induction of human neuronal cells by defined transcription factors. *Nature* **476**, 220–223 (2011).
4. Takahashi, K. & Yamanaka, S. A decade of transcription factor–mediated reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.* **17**, 183–193 (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 (2012).
6. Graf, T. Historical origins of transdifferentiation and reprogramming. *Cell Stem Cell* **9**, 504–516 (2011).
7. Xu, J., Du, Y. & Deng, H. Direct lineage reprogramming: strategies, mechanisms, and applications. *Cell Stem Cell* **16**, 119–134 (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 (2011).
9. Chong, J. J. et al. Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. *Nature* **510**, 273–277 (2014).
10. Mandai, M. et al. Autologous induced stem-cell-derived retinal cells for macular degeneration. *N. Engl. J. Med.* **376**, 1038–1046 (2017).
11. Eiraku, M. et al. Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472**, 51–56 (2011).
12. Lancaster, M. A. et al. Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379 (2013).
13. Takebe, T. et al. Generation of a vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nat. Protoc.* **9**, 396–409 (2014).
14. McCracken, K. W. et al. Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* **516**, 400–404 (2014).
15. Spence, J. R. et al. Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* **470**, 105–109 (2011).
16. Takasato, M. et al. Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. *Nature* **526**, 564–568 (2015).
17. Birey, F. et al. Assembly of functionally integrated human forebrain spheroids. *Nature* **545**, 54–59 (2017).
18. Bagley, J. A., Reumann, D., Bian, S., Lévi-Strauss, J. & Knoblich, J. A. Fused cerebral organoids model interactions between brain regions. *Nat. Methods* **14**, 743–751 (2017).
19. Tang, F. et al. mRNA-seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
This study was the first to carry out whole-transcriptome analysis on single cells.
20. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
21. Klein, A. M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161, 1187-1201. *Cell* **161**, 1187–1201 (2015).

References 20 and 21 were the first studies to take advantage of droplet-based microfluidics to scale up the number of cells for single-cell sequencing.

22. Habib, N. et al. Div-Seq: single-nucleus RNA-seq reveals dynamics of rare adult newborn neurons. *Science* **353**, 925–928 (2016).
23. Shekhar, K. et al. Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308–1323 (2016).
24. Rosenberg, A. B. et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182 (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 (2014).
26. Villani, A. C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
27. Pollen, A. A. et al. Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015).
This single-cell transcriptomic description of the developing human neocortex identified the gene expression phenotype of a cell population (outer radial glia) specifically enriched in the human cortex.
28. Nowakowski, T. J. et al. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (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 (2017).
30. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (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. *Nat. Biotechnol.* **36**, 421–427 (2018).
32. Butler, A. & Satija, R. Integrated analysis of single cell transcriptomic data across conditions, technologies, and species. *bioRxiv* Preprint at https://www.biorxiv.org/content/early/2017/07/18/164889 (2017).
33. Kiselev, V. Yu., Yiu, A. & Hemberg, M. scmap—a tool for unsupervised projection of single cell RNA-seq data. *bioRxiv* Preprint at https://www.biorxiv.org/content/early/2017/11/29/150292 (2017).
34. van den Brink, S. C. et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).
35. Alpert, A., Moore, L. S., Dubovik, T. & Shen-Orr, S. S. Alignment of single-cell trajectories to compare cellular expression dynamics. *Nat. Methods* **15**, 267–270 (2018).
36. Cacchiarelli, D. et al. Aligning single-cell developmental and reprogramming trajectories identifies molecular determinants of reprogramming outcome. *bioRxiv* Preprint at https://www.biorxiv.org/content/early/2017/03/30/122531 (2017).
37. Stegle, O., Teichmann, S. A. & Marioni, J. C. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* **16**, 133–145 (2015).
38. Cahan, P. et al. CellNet: network biology applied to stem cell engineering. *Cell* **158**, 903–915 (2014).
39. Symmons, O. & Raj, A. What's luck got to do with it: single cells, multiple fates, and biological nondeterminism. *Mol. Cell* **62**, 788–802 (2016).
40. Soneson, C. & Robinson, M. D. Bias, robustness and scalability in single-cell differential expression analysis. *Nat. Methods* **15**, 255–261 (2018).
41. Treutlein, B. et al. Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* **534**, 391–395 (2016).
This analysis of direct reprogramming established a precedent to use scRNA-seq to study reprogramming efficiency and revealed that an abundant alternative non-neuronal fate competes with the target neuronal fate.
42. Bargaje, R. et al. Cell population structure prior to bifurcation predicts efficiency of directed differentiation in human induced pluripotent cells. *Proc. Natl. Acad. Sci.* **114**, 2271–2276 (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 (2014).
This paper illustrates that predictions from transcriptomic data can be used to steer differentiation into the desired target fate.
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 (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 (2017).
46. La Manno, G. et al. Molecular diversity of midbrain development in mouse, human, and stem cells. *Cell* **167**, 566–580 (2016).
The authors generate a comprehensive atlas of a developing human tissue and then compare the accuracy and precision of iPSC-derived cells to 'prototypes' in the primary tissue.
47. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
This study demonstrated the ability to infer developmental progression from variation in RNA expression between single cells.
48. Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F. & Theis, F. J. Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* **13**, 845–848 (2016).
49. Briggs, J. A. et al. Mouse embryonic stem cells can differentiate via multiple paths to the same state. *eLife* **6**, e26945 (2017).
50. Clevers, H. Modeling development and disease with organoids. *Cell* **165**, 1586–1597 (2016).
51. Qian, X. et al. Brain-region-specific organoids using mini-bioreactors for modeling ZIKV exposure. *Cell* **165**, 1238–1254 (2016).
52. Camp, J. G. et al. Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proc. Natl. Acad. Sci. USA* **112**, 15672–15677 (2015).
This paper presents the first systematic comparison of expression profiles between cells of primary tissue and the respective organoids on the single-cell level.
53. Camp, J. G. et al. Multilineage communication regulates human liver bud development from pluripotency. *Nature* **546**, 533–538 (2017).
54. Quadrato, G. et al. Cell diversity and network dynamics in photosensitive human brain organoids. *Nature* **545**, 48–53 (2017).
This paper was the first to analyze iPSC-derived cerebral organoids via high-throughput droplet-based single-cell transcriptomics.
55. Stoeckius, M. et al. Cell "hashing" with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *bioRxiv* Preprint at https://www.biorxiv.org/content/early/2017/12/21/237693 (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 (2006).
57. Vierbuchen, T. et al. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).
58. Rackham, O. J. et al. A predictive computational framework for direct reprogramming between human cell types. *Nat. Genet.* **48**, 331–335 (2016).
59. Tsunemoto, R. et al. Diverse reprogramming codes for neuronal identity. *Nature* **557**, 375–380 (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 (2004).
61. Morrison, S. J. & Spradling, A. C. Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell* **132**, 598–611 (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 (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 (2017).
64. Crosetto, N., Bienko, M. & van Oudenaarden, A. Spatially resolved transcriptomics and beyond. *Nat. Rev. Genet.* **16**, 57–66 (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 (2016).
66. Medaglia, C. et al. Spatial reconstruction of immune niches by combining photoactivatable reporters and scRNA-seq. *Science* **358**, 1622–1626 (2017).
67. Halpern, K. B. et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352–356 (2017).
68. Moor, A. E. et al. Spatial reconstruction of single enterocytes uncovers broad zonation along the intestinal villus axis. *bioRxiv* Preprint at https://www.biorxiv.org/content/early/2018/02/07/261529 (2018).
69. Bhatia, S. N. & Ingber, D. E. Microfluidic organs-on-chips. *Nat. Biotechnol.* **32**, 760–772 (2014).
70. Gjorevski, N. et al. Designer matrices for intestinal stem cell and organoid culture. *Nature* **539**, 560–564 (2016).
References 70–72 couple CRISPR–Cas9 genetic manipulation with a single-cell transcriptomic readout to analyze the effect of genetic perturbations on gene expression networks.
71. Wang, X. et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**, eaat5691 (2018).
72. Jaitin, D. A. et al. Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell* **167**, 1883–1896 (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 (2016).
74. Datlinger, P. et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* **14**, 297–301 (2017).
75. Gilbert, L. A. et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).

76. Montalbano, A., Canver, M. C. & Sanjana, N. E. High-throughput approaches to pinpoint function within the noncoding genome. *Mol. Cell* **68**, 44–59 (2017).
77. Haghverdi, L., Buettner, F. & Theis, F. J. Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics* **31**, 2989–2998 (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* Preprint at https://www.biorxiv.org/content/early/2017/09/27/191056 (2017).
79. La Manno, G. et al. RNA velocity in single cells. *bioRxiv* Preprint at https://www.biorxiv.org/content/early/2017/10/19/206052 (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* Preprint at https://www.biorxiv.org/content/early/2017/04/28/127860 (2017).
This study is one of the first examples of lineage-coupled single-cell transcriptomics.
81. Wagner, D. E. et al. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018).
82. Spanjaard, B. et al. Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat. Biotechnol.* **36**, 469–473 (2018).
83. Farrell, J. A. et al. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science* **360**, eaar3131 (2018).
84. Frieda, K. L. et al. Synthetic recording and in situ readout of lineage information in single cells. *Nature* **541**, 107–111 (2017).
85. Smallwood, S. A. et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* **11**, 817–820 (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 (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**, 344ra84 (2016).
88. Trounson, A. & McDonald, C. Stem cell therapies in clinical trials: progress and challenges. *Cell Stem Cell* **17**, 11–22 (2015).
89. Dutta, D., Heo, I. & Clevers, H. Disease modeling in stem cell-derived 3D organoid systems. *Trends Mol. Med.* **23**, 393–410 (2017).
90. Keren-Shaul, H. et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* **169**, 1276–1290 (2017).
This study used scRNA-seq to show that a subpopulation of microglia is responsible for the pathogenicity of a disease with a complex phenotype.
91. Park, J. et al. Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758–763 (2018).
92. Calderon, D. et al. Inferring relevant cell types for complex traits by using single-cell gene expression. *Am. J. Hum. Genet.* **101**, 686–699 (2017).
93. Kang, H. M. et al. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018).
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 (2011).
95. Maude, S. L. et al. Chimeric antigen receptor T cells for sustained remissions in leukemia. *N. Engl. J. Med.* **371**, 1507–1517 (2014).
96. Couzin-Frankel, J. Worries, confusion after cancer trial deaths. *Science* **354**, 1211 (2016).
97. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
98. Wang, Y. et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155–160 (2014).
99. Visvader, J. E. & Lindeman, G. J. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat. Rev. Cancer* **8**, 755–768 (2008).
100. Clevers, H. The cancer stem cell: premises, promises and challenges. *Nat. Med.* **17**, 313–319 (2011).
101. Gupta, P. B. et al. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644 (2011).
102. Batlle, E. & Clevers, H. Cancer stem cells revisited. *Nat. Med.* **23**, 1124–1134 (2017).
103. Boj, S. F. et al. Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**, 324–338 (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 (2015).

Acknowledgements

We thank K. Sekine and T. Nawy for reading the manuscript and for their very thoughtful guidance. Funding for this work was provided by the Max Planck Society (J.G.C., D.W., and B.T.) and the European Research Council (Starting Grant 758877 to B.T.).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to J.G.C. or B.T.