Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Single-cell states versus single-cell atlases — two classes of heterogeneity that differ in meaning and method

Kevin A Janes

Recent advances have created new opportunities to dissect cellular heterogeneity at the omics level. The enthusiasm for deep single-cell profiling has obscured a discussion of different types of heterogeneity and the most-appropriate techniques for studying each type. Here, I distinguish heterogeneity in regulation from heterogeneity in lineage. Snapshots of lineage heterogeneity provide a cell atlas that catalogs cellular diversity within complex tissues. Profiles of regulatory heterogeneity seek to interrogate one lineage deeply to capture an ensemble of single-cell states. Single-cell atlases require molecular signatures from many cells at a throughput afforded by mass cytometry-based, microfluidic-based, and microencapsulation-based methods. Single-cell states are more dependent on time, microenvironment, and low-abundance transcripts, emphasizing *in situ* methods that stress depth of profiling and quantitative accuracy.

Address
Department of Biomedical Engineering, University of Virginia, Charlottesville, VA 22908 USA

Corresponding author: Janes, Kevin A (kjanes@virginia.edu)

Current Opinion in Biotechnology 2016, **39**:120–125

This review comes from a themed issue on **Systems biology**

Edited by **Mark P Styczynski** and **Fabian J Theis**

For a complete overview see the Issue and the Editorial

Available online 1st April 2016

http://dx.doi.org/10.1016/j.copbio.2016.03.015

Introduction

Cellular heterogeneity is the basis for functional diversity in tissues and organs and defines a characteristic response of cell populations to environmental change [1,2]. The concepts of lineage heterogeneity and regulatory heterogeneity have been transformed by an onslaught of new methods that offer rich molecular details at single-cell resolution [3]. Widespread application of these technologies has begun to blur the distinction between cell type and cell state. This can be good for appreciating the tremendous plasticity of stem cells [4,5] or tumor cells [6,7]. However, there are also many compelling questions about regulatory heterogeneity that are separate from lineage heterogeneity [8] and vice versa [9].

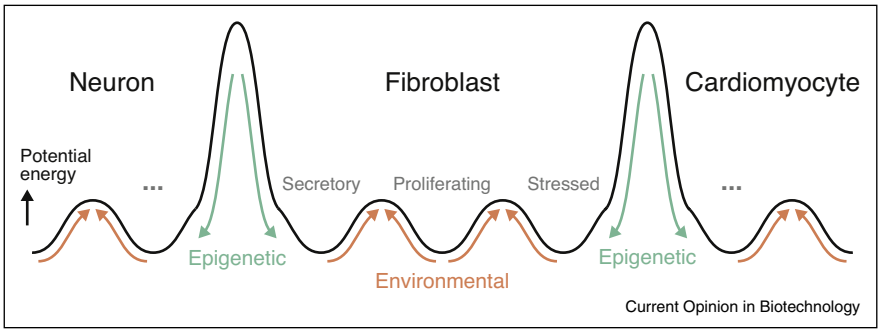
The objective of this review is to organize the single-cell profiling techniques of today according to their relative strengths at capturing regulatory heterogeneity versus lineage heterogeneity. For brevity, I exclude a third source of heterogeneity — transcriptional bursting — given the accumulating evidence of mechanisms that buffer this intrinsic noise in eukaryotes [10–13]. Insightful applications will be emphasized as much as proof-of-concept studies, because the true value of an approach comes with its implementation. Misapplication of a technique to the wrong heterogeneity type can overwhelm some methods and underpower others. Therefore, it is important to think critically about the heterogeneity of interest first and then dedicate time to master the techniques that are most suitable.

Regulatory heterogeneity versus lineage heterogeneity

The difference between regulatory heterogeneity and lineage heterogeneity is more than mere semantics. It speaks to the driving forces of single-cell transitions and the time scale on which they occur. Conceptually, one can merge lineages and states on a graph of cellular 'potential energy' (Figure 1). The energetic barriers between cell types are equivalent to Waddington's landscape, in which epigenetic modifications normally enforce lineage commitment. Within these valleys lie the states that cells of a single lineage can occupy. In contrast to epigenetic barriers of lineage commitment, the hurdles between regulatory states are much smaller. Thus, environmental change is often sufficient to transition cells between regulatory states, the transitions can happen rapidly, and they are frequently reversible. Whereas cell lineages are crucial for development and tissue homeostasis, aberrant cell states are the major cause of adult-onset disease.

Levine and coworkers [14••] recently provided a nice illustration of the difference between single-cell lineage and regulatory state in acute myeloid leukemia (AML). Using mass cytometry to profile surface markers of lineage concurrently with intracellular phosphoproteins indicating regulatory state, the authors tracked leukemic heterogeneity in response to a panel of perturbations. There was a tight coupling between surface markers of myeloid maturation and phospho-signatures in healthy donors, but this relationship broke down in all AML cases examined. The intracellular profiles enabled definition of a regulatory signature for myeloid maturity, which was

Figure 1



Single-cell heterogeneities in lineage (black) and regulatory state (gray). Epigenetic (green) and environmental changes (orange) drive lineage and state transitions respectively. The vertical axis visualizes the extent of gene-regulatory changes between lineages and states as a type of 'potential energy'. Black headings provide representative examples of cell lineage and gray subheadings provide examples of regulatory state.

nonredundant with surface marker lineage and more generalizable across AML patients. This work suggests that regulatory states may catalog diversity within cell and tissue populations more robustly than markers that have historically correlated with lineage.

Heterogeneities in lineage versus regulation bear directly on the single-cell methods that should be employed. In the above mass cytometry application, for example, Levine *et al.* [14••] developed PhenoGraph as an unbiased computational alternative to manual gating, in which surface markers are qualitatively assigned to be present (positive, +) or absent (negative, −). Classically, the best lineage markers are those that separate the positive and negative populations unambiguously. Differences in regulatory state are often much less dramatic, prompting their development of a second algorithm (Statistical Analysis of Response Amplitude, SARA) dedicated to quantitative changes in intracellular signaling [14••]. Similar considerations apply when selecting from the suite of experimental methods now available for single-cell profiling.

Methods for single-cell atlases

How many cell types are there in our lungs [15], or our gut [16], or our brain [17]? We do not really know. Morphology and gross histology underestimate the number of stable cell lineages in any tissue, but how much so remains an open question. The pursuit of undiscovered cell types motivates cell atlasing.

For a molecular profile to be declared characteristic of a new rare cell lineage, it should be undeniably different from the rest of the population. That means looking for outliers. In single-cell transcriptomics, a low outlier of transcript abundance implies 'undetected' [18,19], which is not the same as population-level measurements that can declare a gene absent with confidence [20]. Thus, single-cell atlases rely on transcripts that are sporadically expressed at very high levels compared to the rest of the population. Acknowledging this limitation is powerful for RNA sequencing, because it enables lineage designations to be made with very few total reads per cell [21,22•]. Covariation in biomarkers [23] increases the sensitivity of cell-type discrimination even further. Sequencing of fewer than 9000 transcripts per cell was able to identify rare secretory subtypes among intestinal cells [16], which is remarkable considering that the average cell contains $\sim 10^5$ transcripts [24••].

For cell atlasing, depth requirements may be minimal, but throughput is critical. The major bottlenecks here are single-cell isolation and cell-specific molecular barcoding. Explicit single-cell methods begin with live cells in suspension, invariably requiring extensive enzymatic dissociation of the starting tissue. Tissue dissociation is finicky business [25], and it presumably dislodges cells from their native regulatory state (Figure 1). However, as long as epigenetic barriers hold strong, the thinking is that enough lineage-specific information will be retained in the single-cell molecular profile. Grün and coworkers [16] cleverly work within this limitation by first culturing intestinal stem cells as 3D organoids, enabling faster isolation of single cells compared to primary tissue. Other techniques have long been available to isolate single cells from snap-frozen tissue [26], but single-cell library preparations have not been reported in this format.

The barcoding bottleneck was recently surmounted by three groups [27••,28••,29••], who independently devised microencapsulation-based workflows for cell lysis, RNA capture, and library preparation. The essence of these methods is to isolate a single cell together with a single microsphere containing custom oligonucleotides for RNA capture. The oligonucleotides contain (from $5' \rightarrow 3'$) a universal primer or promoter, a microsphere-specific barcode indicating cell identity, a library of unique molecular identifiers [30], and a common oligo(dT) for binding