DEVELOPMENT

Integrating lineage tracing and single-cell analysis

LARRY unveils transcriptomic changes leading to fate determination in hematopoiesis.

Single-cell genomic profiling and lineage tracing are two categories of powerful tools in developmental biologists' arsenal. Despite their popularity, their weaknesses are as well known as their strengths. "Lineage tracing establishes dynamic relationships over long time intervals. Their drawback is that the specificity by which the state of initially labeled cells can be assessed is limited," says Allon M. Klein, of Harvard Medical School. "Single-cell genomic profiling takes a complementary approach: it can resolve differences between cells with much detail, and do so without assumptions about which features matter. However, these measurements kill cells, so they alone can't be used to relate early states of cells to their future behavior." Thus, a strategy that integrates both technologies has the potential to provide a more comprehensive picture of developmental processes, in terms of cell states and associated dynamics.

To this aim, Klein and colleagues developed lineage and RNA recovery (LARRY), which uses a barcoding system that can be detected by single-cell RNA sequencing (scRNA-seq) for clonal labeling. Each barcode is a random 28-mer in the 3' untranslated region of an enhanced green fluorescent protein transgene delivered by a lentiviral vector. To scale up, libraries of high complexity ($\sim 0.5 \times 10^6$ barcodes) were constructed, allowing for 5,000 cells to be labeled in an experiment. Other optimizations were also needed, as noted by Klein. "Care must be taken to design experiments where the number of clones is optimal—too few, and you lose statistical power; too many, and you start sampling many clones by just one cell, which renders those clonal data useless. A second consideration we had was how to ensure that we could analyze very short-term dynamics. This meant that one needs very high cell-recovery efficiency to analyze small clones."

Klein and colleagues applied LARRY to study hematopoiesis, the process underlying blood regeneration in the bone marrow, where hematopoietic stem and progenitor cells give rise to different lineages of blood cells. Mouse progenitor cells were barcoded with LARRY and underwent
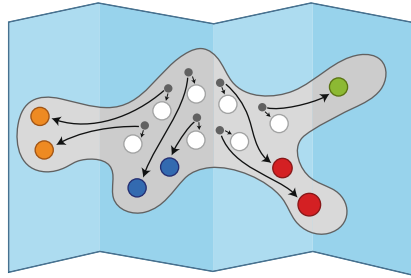
Schematic of the lineage map detected by LARRY.

division. Then half of the progeny cells were harvested for scRNA-seq to define the 'early state'. The other half were allowed to undergo further rounds of division and differentiation through in vitro culture or in vivo transplantation, and were harvested for scRNA-seq to define the 'late stage'. The identities of barcodes established lineage relationships among cells. Leveraging the fine resolution of scRNA-seq, LARRY built a rich "map of clonal fate on a continuous transcriptional landscape," which is different from the classical view of hematopoiesis, as described by a set of discrete states and their transitions.

Further analysis unveiled more details about fate decisions in hematopoiesis. "One surprise we had was that cells appeared to be strongly committed to one fate or another well before they reached a 'branchpoint' toward that fate on the single-cell-generated landscape. The implications for hematopoiesis are that cells appear to make decisions much earlier than we thought, and much earlier than you might guess by looking at scRNA-seq data alone. This tells us where to focus next in studying decision-making." Klein and colleagues also tried to answer "whether scRNA-seq captures the state of a cell well enough to predict behavior." By comparing transcriptomes of sister cells under different experimental conditions, they showed evidence of the existence of 'hidden variables' that are not measured in scRNA-seq data but influence cell fates. While the exact nature and function of such variables await future investigation, it is clear that better sampling of RNA or detection of other cellular

components is needed for predicting the fate choice of a cell.

Beyond experimentalists, computational biologists might also find these data of interest. Development of computational methods for analyzing scRNA-seq data is a thriving area, with goals such as inferring pseudotime trajectories and predicting cell-fate choice. However, experimental data of 'ground truth' for testing and benchmarking these algorithms are generally limited. Analysis performed by Klein and colleagues has shown that although pseudotime estimated by algorithms often makes reasonable sense, fate-choice predictions by the current computational methods are not always reliable. "We've already had a couple of groups ask us for the data. Because LARRY approximately relates the early state of a cell to its future fate, this data is excellent for training or testing machine-learning algorithms that hope to predict future dynamics from the state of a cell," notes Klein.

His team looks forward to future development of LARRY, in terms of both the technology and its applications. "One exciting technological direction is to repurpose LARRY for other single-cell genomic readouts," says Klein. "We now have the opportunity to lineage-trace based on chromatin state, composite proteomic state, et cetera." Regarding applications to consider, he says, "For us, an exciting application is to combine LARRY with perturbation analysis. A simple design is to perturb one sister cell while measuring the state of the other. Through this, LARRY offers the possibility of identifying how perturbations act in a state-dependent manner. This may hold a key to understanding the action of pleiotropic signaling factors, for example."

Lin Tang

Research paper
Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D. & Klein, A. M. Lineage tracing on transcriptional landscapes links state to fate during differentiation. Science 367, eaaw3381 (2020).