Tracing Cell Lineages from Single-Cell Data using Markov Affinity Estimation

Li, Lauren (School: Westview High School)

Lineage tracing provides fundamental insights into the developmental trajectories of individual cells of complex organisms. Genetic labeling approaches, such as naturally-occurring mitochondrial mutations and CRISPR-based barcoding, provide powerful means for tracing the development of individual cells originating from a single zygote. However, technical noise introduced by single-cell RNA-sequencing, coupled with limitations in cellular barcoding, creates challenges for accurate developmental lineage profiling at single-cell resolution. Thus, we developed PRISM (PRogeny Identification from Single-cell Mutations), a computational framework for reconstructing developmental lineages using a Markov affinity-based approach, which recovers lost cell-cell relationships from high-dimensional single-cell data. PRISM analyzes the genetic affinity of cells to learn the underlying developmental structure, with inherent algorithmic robustness against error-prone feature measurements, resulting in accurate reconstruction of lineage trees. From a gold-standard in vitro barcoding dataset, PRISM yields >96% accuracy in reconstructing experimental lineages, outperforming canonical Euclidean distance-based approaches. Analysis of an in vivo dataset of cells in a human colorectal cancer tumor via PRISM revealed a clonal subpopulation of LGR5+ and MKI67+ malignant cells, indicating this subpopulation of cells proliferated from a specific cell-of-origin. Furthermore, PRISM reconstructs broad developmental lineages in zebrafish using CRISPR-Cas9 barcoding. Overall, PRISM is scalable and robust, capable of handling data from multiple lineage tracing systems in multiple complex organisms and yielding new insights into the development of healthy and diseased cells.

Awards Won:
Second Award of $1,500