

Monday Minisymposium

Organisers: MacLean & Rockne

Laleh Haghverdi

Max Delbrück Center for Molecular Medicine, Berlin

"Cell lineage trajectories and pseudotime reconstruction from single-cell transcriptomics"

The reconstruction of cell lineage trajectories from single-cell transcriptomics allows us to resolve temporal expression dynamics of several genes from snapshot collected data. The temporal order of gene activities in return, can provide new insights into the gene regulatory networks governing cell differentiation. We adapted a dimension-reduction technique called "diffusion maps" for the analysis and reconstruction of cell differentiation trajectories. We perceived cell differentiation as a diffusion-like process, where cells are gradually and to some extent stochastically changing their gene expression profiles as they proceed to more differentiated molecular states. Thus, cell differentiation dynamics can (in discrete form) be described by: $p(t) = p(t-1) * T$. That is, the probability distribution at the position of each sampled cell at time t , $p(t)$ is given by the probability distributions at time $(t-1)$ multiplied by the cells' pairwise transition probabilities matrix T . Thus, $p(t)$ and $p(t-1)$ are vectors of length N (number of cells) and T is an N by N matrix which accommodates: 1) transition probabilities which are purely based on geometrical distances between the cells, 2) directional transition probabilities towards more differentiated cell states and 3) source/sink probabilities for accounting for cell's proliferation/death rates. Considering purely geometrical transition probabilities, T_{geom} is a row-normalised positive-definite matrix which characterises the manifold on which cell differentiation is taking place in the high-dimensional gene expression space. The complete process (i.e., including directional and source/sink diffusion terms) thus, describes different dynamics taking place on that same manifold. Because of this resemblance to a diffusion process, we introduced an adaptation of diffusion maps for data embedding and differentiation trajectories reconstruction from single-cell data. In a simple form, data embedding by diffusion maps is given by the first eigenvectors of T_{geom} corresponding to its largest eigenvalues. We extended the diffusion map's framework to define "pseudotime" quantifying the differentiation stage of each cell based on on-manifold diffusion distances from the root (pluripotent) cell state. Furthermore, we showed that differentiation branching events can be identified by application of the triangle inequality to the on-manifold distances. The diffusion metrics with new approaches for inclusion of directional and source/sink probabilities (inferred from biologically known differentiation terminal points, spliced/unspliced mRNA comparisons, time-lapse measurements for a few genes) has been used in several trajectory reconstruction methods on exceedingly complex manifolds which I will briefly discuss in my presentation.

Adam MacLean

University of Southern California

"Inferring the structure and parameters of cell lineage models from single-cell data"

Since single-cell RNA sequencing technologies have become widespread, great efforts have been made to develop appropriate statistical methods to learn biological features from high dimensional data. Lagging behind are efforts to combine dynamical systems models with single-cell data. The existence of transient or intermediate cell states – corresponding to shallow attractor states on a potential landscape – further confounds the development of simple models. Here we present new approaches to address these challenges. We propose methods to infer both model structure and model parameters of cell lineage models using single-cell RNA-sequencing data. We fit model structure by assuming a core set of cell population dynamic processes, i.e. a general system described by $\dot{x}(t) = (\alpha f(x) - \delta)x(t)$, where $x(t)$ is a vector of cell states, α and δ are cell state-specific proliferation and death rates, and $f(x)$ is a feedback function. We use methods for trajectory and cell-cell communication inference to infer lineage relationships and feedback interactions. Taking advantage of the information present in the pseudotemporal ordering of cells, we perform Bayesian parameter inference: that is we fit the model to "pseudo-dynamics" to constrain the parameters. With application to hematopoiesis, we demonstrate the ability to select between models and predict future cell differentiation dynamics. We will discuss how this model inference framework can also be applied to gene regulatory network dynamics, enabling not only the simulation of single-cell data-derived models, but also the characterization of their stable steady states and bifurcations.

Caleb Weinreb

Harvard University

"Limits on dynamic inference from single-cell RNA-seq and the case for additional measurements"

For decades, biologists have dreamed of building dynamical systems models of gene expression that could be used to understand and predict cell state. High-throughput single-cell analysis has brought the dream closer. It is now possible to sample cell states in high-dimension and to map trajectories through continuous gene expression space. Using prior knowledge or measurements of RNA maturity, the general direction of cell progression along these trajectories can be inferred. Yet it remains unclear to what extent single-cell RNA-seq encodes a unique dynamical system over cell state. Here, I will discuss two key obstacles to inferring cell state dynamics from single-cell RNA seq data, and how additional types single-cell measurements might overcome them. First, applying the principle of population reveals the importance of measuring cell proliferation and death rates for accurately inferring cell dynamics. Second, clonal tracing using DNA barcoding reveals that RNA-seq alone cannot distinguish cells with different cell-autonomous fate biases, emphasizing the need to measure other variables such as chromatin state.

Heyrim Cho

University of California at Riverside

"Modeling cell state dynamics of Hematopoiesis from single-cell gene sequencing data"

Recent advances in single-cell gene sequencing data and high-dimensional data analysis techniques are bringing new opportunities in modeling biological systems. In this talk, I will discuss different approaches to develop mathematical models from high-dimensional single-cell data. In particular, single-cell RNA sequencing data is challenging due to its high-dimensionality, and dimension reduction techniques are essential in finding the trajectories of cell states in the reduced differentiation space. We develop models using differential equations that describe the cell differentiation as directed and random movement on the abstracted graph or on the multi-dimensional reduced space. Normal hematopoiesis differentiation and abnormal processes of acute myeloid leukemia (AML) progression are simulated. We show that the model can predict the emergence of cells in novel intermediate states of differentiation consistent with immunophenotypic characterizations of AML, and compare the pros and cons of the models on the graph and on the multi-dimensional reduced space.