Yale SCHOOL OF PUBLIC HEALTH

Biostatistics

Exploring Space and *Time* for Identifying Gene Interactions Using Single-cell Transcriptomics

Atul Deshpande, PhD
Postdoctoral Researcher
Division of Biostatistics and Bioinformatics
Johns Hopkins University

12:00 Noon Eastern time, Tuesday, October 5, 2021

ABSTRACT

Advances in single-cell transcriptomics enable measuring the gene expression of individual cells, thus allowing these cells to be ordered along a trajectory of a dynamic biological process. Many trajectory inference algorithms assign `pseudotimes' to each cell, representing the progress along the biological process. Ordering the expression data according to such pseudotimes can be valuable for understanding the underlying regulator-gene interactions in a biological process, such as differentiation. However, the distribution of cells sampled along a transitional process, and hence that of the pseudotimes assigned to them, is not uniform. This prevents the use of many standard mathematical methods for analyzing the ordered gene expression states. In the first part of my talk, I will describe Single-cell Inference of Networks from Granger Ensembles (SINGE), an algorithm for gene regulatory network inference from single-cell gene expression data. Given ordered single-cell data, SINGE uses kernel-based Granger Causality regression, which smooths the irregular pseudotimes and missing expression values. It then aggregates the predictions from an ensemble of regression analyses with a modified Borda count to compile a ranked list of candidate interactions between transcriptional regulators and their target genes. I will also discuss considerations for using pseudotimes obtained from trajectory inference methods for further analyses.

In the second part of the talk, I will briefly describe my current research focus on using spatial transcriptomics to identify gene interactions. Recent advances in sequencing technology have resulted in the development of spatial transcriptomics, which enable in situ tissue profiling of gene expression data at near single-cell resolution. We use CoGAPS, a Bayesian non-negative matrix factorization technique to learn latent biological patterns from spatial transcriptomics. We identify hotspots of pattern activity and regions of interaction using the learned patterns from CoGAPS, subsequently identifying genes corresponding to pattern interactions.