Statistics Seminar

Applications of statistical and machine learning methods in spatial transcriptomics

Mingyao Li, University of Pennsylvania

Abstract: Recent developments in spatial transcriptomics technologies have enabled scientists to get an integrated understanding of cells in their morphological context. Applications of these technologies in diverse tissues and diseases have transformed our views of transcriptional complexity. Most published studies utilized tools developed for single-cell RNA-seq for data analysis. However, spatial transcriptomics data exhibit different properties from single-cell RNA-seq. To take full advantage of the added dimension on spatial location information in such data, new methods that are tailored for spatial transcriptomics are needed. Additionally, spatial transcriptomics data often have companion high-resolution histology information available. Incorporating histological features in gene expression analysis is an underexplored area. In this talk, I will present several statistical and machine learning methods that we recently developed for spatial transcriptomics data analysis. Specifically, I will discuss how to 1) perform spatial clustering, 2) infer cell type distribution by deconvolution, 3) generate super-resolution gene expression, 4) integrate gene expression and histology for tissue annotation, 5) recover spatial location information for single cells, and 6) predict gene expression from histology. I will show applications of these methods in spatial transcriptomics data generated from brain and cancer tissues.

Bio: Dr. Li is a Professor of Biostatistics at the Department of Biostatistics, Epidemiology and Informatics University of Pennsylvania Perelman School of Medicine.
Dr. Li received her PhD in Biostatistics from the University of Michigan in 2005. She was trained as a statistical geneticist, but since she joined the faculty at the University of Pennsylvania in 2006, she has gradually transitioned her research from traditional statistical genetics to statistical genomics with the goal of having a deeper understanding of the molecular mechanism of human disease. The central theme of her current research is to use statistical and machine learning methods to understand cellular heterogeneity in human-disease-relevant tissues, to characterize gene expression diversity across cell types, to study the patterns of cell state transition and crosstalk of various cells using data generated from single-cell and spatial transcriptomics studies, and to translate these findings into the clinics. In addition to methods development, she is also interested in collaborating with researchers seeking to identify complex disease susceptibility genes and acting cell types. At UPenn, she serves as the Director of Biostatistics for the Gene Therapy Program, where she advises biostatistics and bioinformatics analysis for various gene therapy studies. She also chairs the Graduate Program in Biostatistics. She is an elected member of the International Statistical Institute and a Fellow of the American Statistical Association.

Email: mingyao@pennmedicine.upenn.edu
Website: https://transgen.med.upenn.edu