COLLOQUIUM

Hyun Min Kang
University of Michigan

Mixture Models for Sequence Contamination and Single Cell Transcriptomics

Tuesday, November 22, 2016
10:20 a.m. - 11:10 am
Refreshments 10:00 am
C405 Wells Hall

Abstract

It is surprisingly common that DNA and RNA sequence reads are contaminated by unexpected samples during various steps of the high throughput sequencing. Contaminated sequenced reads, if neglected, may result in substantial false positives and reduced power in downstream analyses. In this talk, I will describe a series of mixture model methods to accurately detect, estimate, and correct for within-species sample contamination in various settings of DNA and RNA sequencing studies. I will demonstrate how substantially these methods contribute to improving the quality of genetic analyses with real-world examples.

In addition, I will also demonstrate how these models can be extended to design powerful single-cell RNA-seq studies. Our methods enable efficient sample multiplexing for droplet-based single-cell transcriptome profiling, substantially reducing costs and batch effects, and increasing the sample size and statistical power in genetic studies of single-cell transcriptomes.

To request an interpreter or other accommodations for people with disabilities, please call the Department of Statistics and Probability at 517-355-9589.