The Department of Biomedical Data Science

Presents

**Hongzhe Lee**
Perelman Professor of Biostatistics, Epidemiology and Informatics,
University of Pennsylvania School of Medicine

*“Interrogating the Gut Microbiome: Estimation of Bacterial Growth Rates and Prediction of Biosynthetic Gene Clusters”*
Thursday, April 15, 2021
2:30-3:50 PM PST
[https://stanford.zoom.us/j/95469651140?pwd=RXhYM21Fc1ISVWdxQ3I2emlyVndHQT09](https://stanford.zoom.us/j/95469651140?pwd=RXhYM21Fc1ISVWdxQ3I2emlyVndHQT09)
Passcode: 206507

**Abstract**
The gut microbiome plays an important role in maintenance of human health. High-throughput shotgun metagenomic sequencing of a large set of samples provides an important tool to interrogate the gut microbiome. Besides providing footprints of taxonomic community composition and genes, these data can be further explored to study the bacterial growth dynamics and metabolic potentials via generation of small molecules and secondary metabolites. In this talk, Dr. Lee will present several computational and statistical methods for estimating the bacterial growth rate for metagenome-assembled genomes (MAGs) and for predicting all biosynthetic gene clusters (BGCs) in bacterial genomes. The key statistical and computational tools used include optimal permutation recovery based on low-rank matrix projection and improved LSTM deep learning methods to improve prediction of BGCs. He will demonstrate the application of these methods using several ongoing microbiome studies of inflammatory bowel disease at the University of Pennsylvania.

**Brief Bio**
Hongzhe Li is Perelman Professor at the University of Pennsylvania Perelman School of Medicine. He is Vice Chair of Research Integration and Director of Center for Statistics in Big Data in the Department of Biostatistics, Epidemiology and Informatics. He is an elected fellow of ASA, IMS and AAAS. His research interest includes statistical genetics, metagenomics and high dimensional statistics.

**Suggested Readings**
- [Quantifying and comparing bacterial growth dynamics in multiple metagenomic samples](#)
- [Optimal Permutation Recovery in Permuted Monotone Matrix Model](#)