

# Myra Samuels Memorial Lecture 2020

Friday, March 6, 2020

10:30 – 11:30 a.m.

Lawson (LWSN) 1142

\*There will be a reception for Dr. Huang in HAAS 111 at 10:00 a.m.

## Dr. Haiyan Huang

Professor of Statistics and Director of the Center for Computational Biology  
University of California, Berkeley

### ***GeneFishing: a method to reconstruct context-specific portraits of biological processes and its application to cholesterol metabolism***

*Abstract:* Rapid advances in genomic technologies have led to a wealth of diverse data, from which novel discoveries can be gleaned through the application of robust statistical and computational methods. Here we describe GeneFishing, a computational approach to reconstruct context-specific portraits of biological processes by leveraging gene-gene co-expression information. GeneFishing incorporates multiple high-dimensional statistical ideas, including dimensionality reduction, clustering, subsampling and results aggregation, to produce robust results. To illustrate the power of our method, we applied it using 21 genes involved in cholesterol metabolism as “bait”, to “fish out” (or identify) genes not previously identified as being connected to cholesterol metabolism. Using simulation and real datasets, we found the results obtained through GeneFishing were more interesting for our study than those provided by related gene-prioritization methods. In particular, application of GeneFishing to the GTEx liver RNAseq data not only re-identified many known cholesterol-related genes, but also pointed to glyoxalase I (GLO1) as a novel gene implicated in cholesterol metabolism. In a follow-up experiment, we found that GLO1 knock-down in human hepatoma cell lines increased levels of cellular cholesterol ester, validating a role for GLO1 in cholesterol metabolism. In addition, we performed pan-tissue analysis by applying GeneFishing on various tissues and identified many potential tissue-specific cholesterol metabolism related genes. GeneFishing appears to be a powerful tool for identifying novel related components of complex biological systems and may be employed across a wide range of applications.