



# THE UNIVERSITY OF CHICAGO

Department of Statistics

## STATISTICS COLLOQUIUM

---

XIN HE

Department of Human Genetics  
University of Chicago

### Integrative Statistical Methods for Mapping Disease Genes

MONDAY, February 16, 2015, at 4:00 PM  
Eckhart 133, 5734 S. University Avenue  
*Refreshments following the seminar in Eckhart 110*

#### ABSTRACT

Biology is increasingly becoming a "data science": hundreds of thousands of human genomes are being sequenced; large amount of gene expression, protein-DNA interaction, and other types of genomic data are available. The key challenge is to extract "meaning" from data, to benefit our understanding of human diseases. In this talk, I will describe my recent work on identifying risk genes for complex diseases by novel and integrative methods. First, I will show that integrating multiple types of genetic variants leads to a powerful strategy of genetic mapping. Each person inherits mutations from parents, some of which may predispose the person to certain diseases. Meanwhile, new mutations may occur spontaneously during the reproductive process, and if disrupting key genes, these de novo mutations may increase risks of disease. We developed a Bayesian model that effectively combines data from de novo mutations, inherited variants in families, and standing variants in the population (identified with case-control studies). This approach greatly increases the power of gene discovery and predicts promising genes for autism. In the second part of my talk, I will describe a method we developed, named Sherlock, that jointly analyzes expression QTL and data from genome-wide association studies (GWAS). This method allows us to effectively combine many weak signals in GWAS to identify disease susceptibility genes. Because many such signals are linked to expression of a gene in trans, Sherlock is able to detect completely new genes from GWAS, and we made promising discoveries in several different diseases.