



RNA Innovation Seminar
Monday, January 8th at 3:00pm
**ABC Seminar rooms, Biomedical Research
Science Building (BSRB), 109 Zina Pitcher**

[Xiang Zhou, Ph.D.](#)
Assistant Professor
Department of Biostatistics

“MACAU: Differential Expression Analysis and Expression Heritability Estimation for RNAseq Studies”

Keywords: differential expression; heritability; RNAseq; Poisson mixed model

Abstract: Identifying differentially expressed (DE) genes from RNA sequencing (RNAseq) studies is among the most common analyses in genomics. However, RNAseq DE analysis presents several statistical and computational challenges, including over-dispersed read counts and, in some settings, sample non-independence. Previous count-based methods rely on simple hierarchical Poisson models (e.g., negative binomial) to model independent over-dispersion, but do not account for sample non-independence due to relatedness, population structure and/or hidden confounders. Here, we present a Poisson mixed model with two random effects terms that account for both independent over-dispersion and sample non-independence. We also develop a scalable sampling-based inference algorithm using a latent variable representation of the Poisson distribution. With simulations, we show that our method properly controls for type I error and is generally more powerful than other widely used approaches, except in small samples ($n < 15$) with other unfavorable properties (e.g., small effect sizes). We also apply our method to three real data sets that contain related individuals, population stratification, or hidden confounders. Our results show that our method increases power in all three data compared to other approaches, though the power gain is smallest in the smallest sample ($n=6$). Finally, we also extend our methods for gene expression heritability estimation with sequencing data. Our method is implemented in MACAU, freely available at www.xzlab.org/software.html.