

数学与系统科学研究院

计算数学所学术报告

报告人: **Prof. Zhaohui Qin**

(*Department of Biostatistics and Bioinformatics*

*Emory University, USA*)

报告题目:

**Statistical Modeling of RNA-seq  
Data**

邀请人: 唐贻发研究员

报告时间: **2011年1月6日(周四)**

**上午 10:00**

报告地点: **科技综合楼三层 311**

**计算数学所报告厅**

## **Abstract:**

RNA sequencing (RNA-seq) is a powerful new technology for mapping and quantifying transcriptomes using ultra high-throughput next generation sequencing technologies. Using deep sequencing, gene expression levels can be quantified thus providing a digital measure of the presence and prevalence of all transcripts including novel ones. Although extremely promising, the massive amounts of data that are generated by RNA-seq, substantial biases, and uncertainty in short read alignment pose daunting challenges for data analysis. In particular, large base-specific variations and between-base correlations make naive approaches, such as averaging to normalizing RNA-seq data and quantifying gene expressions, ineffective. We propose to develop Poisson mixed effects models to characterize RNA-seq data. These models will accommodate the biases, variations, and correlations present in RNA-seq data so as to accurately estimate gene expression levels and to facilitate gene expression comparison and novel transcript structure or activities discovery.

欢迎大家参加!