

学术报告会

时间：2013年5月14日(周二)14:00-15:30

地点：电院群楼2-410会议室

Studying the Expression of Alternative Splicing Genes with RNA-Seq Data



Prof. Xuegong Zhang

Tsinghua University

Abstract:

Studying the expression of genes at mRNA level is a basic step for many biological studies. The fast advancing next generation sequencing technology has opened a new page for measuring RNA abundances with deep sequencing technologies or RNA-Seq. RNA-Seq has revealed many new observations on the characteristics of mRNA expression, especially on the wide existence of alternative splicing, which makes the estimation of gene expression values and identification of differentially expressed genes not a straightforward task. We have conducted a series of work on the estimation of expression of alternative isoforms and of the whole gene, and comparing differential expression and differential splicing of genes between samples or groups of samples. One basic challenge is the modeling RNA-Seq data that have non-uniform distribution of reads on the transcripts. In this talk, I'll report our work on using non-uniform read distribution models to improve isoform expression inference. We developed a non-parametric method to model and correct for the non-uniform distribution of reads. Experiments on both simulated and real data showed that the method can estimate the expression of isoforms and of the whole genes more accurately, and can identify major isoforms with higher fidelity. I'll also share our latest study on detecting differentially spliced genes between two groups of samples, and brief on a few other bioinformatics and systems biology projects at the Bioinformatics Division of Tsinghua National Laboratory for Information Science and Technology.

Biography:

Prof. Xuegong Zhang earned his BS degree in Industrial Automation in 1989 and Ph.D. degree in Pattern Recognition and Intelligent Systems in 1994, both from Tsinghua University. He joined the faculty of Tsinghua University in 1994, where he is a Professor of Pattern Recognition and Bioinformatics now. Dr. Zhang worked at Harvard School of Public Health as a visiting scientist on computational biology in 2001-2002 and in 2006. Currently he is the Director of the Bioinformatics Division, Tsinghua National Laboratory for Information Science and Technology. His research interests include machine learning, biological data mining especially for microarray and deep sequencing data, gene expression and regulation, alternative splicing, metagenomics methods and applications, etc..