

Pei Fen Kuan, PhD

Department of Biostatistics & Medical Informatics
University of Wisconsin-Madison

A Statistical Framework for the Analysis of ChIP-Seq Data

Abstract:

Chromatin immunoprecipitation followed by direct sequencing (ChIP-Seq) has revolutionized the experiments in profiling DNA-protein interactions and chromatin remodeling patterns. Although this technology offers promising results for surveying large genomes at higher resolution, it is not free of sequencing and other sources of biases. Despite this, most of the existing tools do not consider such biases. We carefully study sources of bias in the underlying data generating process of ChIP-Seq technology by utilizing sequenced naked DNA (non-cross-linked, deproteinized DNA) and develop a model that captures the background signal in the ChIP-Seq data. We then proposed mixture models for analyzing ChIP-Seq data. Our modeling framework incorporates the variability in both the mappability and GC-content of regions on the genome and sequencing depths of the samples. We show that our model fits very well on real data and provides a fast model-based approach for ChIP-Seq data analysis.

Tuesday, September 29th, 2009

4:00 pm

Biotechnology Center Auditorium

425 Henry Mall