THE UNIVERSITY OF CHICAGO

Department of Statistics

MASTER'S THESIS PRESENTATION

XINYUE LI

Department of Statistics
The University of Chicago

Inference of Genetic Similarity between Tissues from RNA-seq Data

THURSDAY, May 15, 2014, at 10:00 AM
Eckhart 117, 5734 S. University Avenue

ABSTRACT

With a growing number of genetic datasets from human tissues, many research studies have been conducted to understand tissue similarities and differences. However, no universal metric for quantifying the genetic similarity between human tissues has been developed, due to the diversity in the types of genetic datasets spanning from microarray data to eQTL mapping data as well as to the difficulties in understanding the delicate biological features of human tissues. Using the RNA-seq dataset from the GTEx project, we evaluate some of the traditional metrics and develop a new method for inferring tissue similarity based on RNA-seq data. Traditional methods often use the RNA expression levels of each tissue sample across all gene transcripts to calculate Pearson's correlation and determine tissue similarity. In contrast, we compare the RNA expression levels of each gene transcript across tissue samples and use the distribution of gene-level correlations to infer tissue similarity. We also discuss ways to combine different statistics into measures of similarity by weighting them according to the application of interest.