ABSTRACT

Transcriptome data provides key information about molecular mechanism for phenotypic diversity. Advances in technology have made transcriptome data available in improved quality and quantity, calling for new statistical methods that can account for large data size, complex dependence structure, and technical artifacts. This dissertation proposes statistical methods that tackle those challenges in transcriptome data analysis while addressing important biological questions.

Chapters 2, 3, and 4 focus on the analysis of African American gene expression levels. African American samples' genetic ancestry is investigated regarding its relationship with their gene expression levels. Chapter 5 introduces a statistical method for data sets that became more recently available --- single cell transcriptome. Chapter 5 provides a comprehensive analysis tool for scRNA-seq data from 10X protocol by modeling the noise structure. Although the proposed methods are developed for the specific purpose of analyzing gene expression level data, some of them can be potentially applied to diverse fields. For example, Chapter 2 develops a multivariate Bayesian variable selection tool that can account for data sets with random missing values.

Chapter 4 develops a covariance analysis tool that expands traditional heteroskedasticity analysis to dynamically varying covariance analysis. These methods focus on accounting for inter-tissue or inter-gene correlations, so they can be applied to correlated data from other fields. Most of the methods have been implemented as open-source softwares to promote their applicability. We
believe these methods contribute not only to future research in molecular biology but also to the field of large and complex, modern data analysis.

For information about building access for persons with disabilities, please contact Keisha Prowoznik at 773.702-0541 or send an email to kprowoznik@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following web site: https://lists.uchicago.edu/web/subscribe/statseminars.