Statistical Methods for Expression Quantitative Trait Loci (eQTL) Mapping using RNA-seq Data

Wei Sun

Department of Biostatistics
University of North Carolina, Chapel Hill

Abstract

The study of gene expression Quantitative Trait Loci (eQTL) is attracting great interests from scientists who study the genetic basis of gene expression or the biological processes that bridge genetic variations and complex diseases. Several statistical methods have been developed for eQTL mapping while gene expression is measured by microarray. Recently, high-throughput sequencing techniques have become popular approaches to measure RNA abundance (i.e., RNA-seq), and it is expected to replace microarray in the near future. Using RNA-seq, expression level of a gene can be measured by the Total Read Count (TReC). In contrast to the expression levels measured by microarray probe intensities, RNA-seq measurements are more accurate and more sensitive. In addition, RNA-seq also provides the information of allele specific expression (ASE), which is not available from microarray. If properly used, ASE information can improve the power and precision of eQTL mapping. In this paper, we develop a statistical method to map eQTL using RNA-seq data by combining the TReC measurement and ASE measurement. Our new method, which is named as TReCASE, shows improved performance than traditional methods in both simulation and real data studies.