Microarray Data Visualization and Preprocessing Using R and Bioconductor

Authors : Ruchi Yadav, Shivani Pandey, Prachi Srivastava

Abstract : Microarrays provide a rich source of data on the molecular working of cells. Each microarray reports on the abundance of tens of thousands of mRNAs. Virtually every human disease is being studied using microarrays with the hope of finding the molecular mechanisms of disease. Bioinformatics analysis plays an important part of processing the information embedded in large-scale expression profiling studies and for laying the foundation for biological interpretation. A basic, yet challenging task in the analysis of microarray gene expression data is the identification of changes in gene expression that are associated with particular biological conditions. Careful statistical design and analysis are essential to improve the efficiency and reliability of microarray experiments throughout the data acquisition and analysis process. One of the most popular platforms for microarray analysis is Bioconductor, an open source and open development software project based on the R programming language. This paper describes specific procedures for conducting quality assessment, visualization and preprocessing of Affymetrix Gene Chip and also details the different bioconductor packages used to analyze affymetrix microarray data and describe the analysis and outcome of each plots.