Bioconductor: A system for statistical analysis in functional genomics

Robert Gentleman
Harvard University and
The Dana Farber Cancer Inst., USA
rgentleman@hsph.harvard.edu

Abstract

I will discuss an open source system for the analysis of geneomic data. The system is centered around R (www.r-project.org) and provides direct connections to specialty image storage libraries, hierarchical (or object-oriented) databases for the storage of the relevant data. It combines, in a coherent method for organizing and accessing quantitative data, phenotypic data and experimental data. The system provides access to many statistical routines for both computation and visualization.

Practical examples of how the system can be used to rapidly develop packages to perform specific tasks such as genefiltering, ROC curves, effective discrimination of gene expression patterns between classes of phenotypes. If time permits I will also consider some new visualization tools that are currently in development.