

MA570: Data Analysis for high-throughput genomics technologies

Course Outline

- Introduction to Linux & the R Statistical Computing Environment
- The BioConductor project
- Microarrays: Experimental design and analysis issues
- Analysis of microarray data using BioConductor, including
 - Pre-processing & quality control
 - Differential expression analysis
 - Geneset enrichment analysis
 - Case studies
- Next-generation sequencing technologies and their applications
- RNA-seq: Gene expression analysis using next-generation sequencing, including
 - Quality control using the fastq pipeline
 - Mapping short-reads to reference genomes
 - Gene expression metrics and caveats
 - Case studies
- Chromatin Immuno Precipitation and next-generation sequencing (ChIP-seq), including
 - Design and analysis of ChIP-seq experiments
 - Peak-finding
 - Case studies (including the Encode project)