hms | hsdm
office for postdoctoral fellows
25 Shattuck Street | Gordon Hall 006 | Boston, MA 02115 | 617.432.6735 | postdoc_office@hms.harvard.edu | postdoc.hms.harvard.edu

Bioinformatics Training for HMS/HSDM Postdoctoral Fellows
(in partnership with the Countway Library and the Harvard Catalyst)

Introduction to Bioinformatics (1.5h)

Description: This workshop will give overview of basic concepts and fundamentals underlying modern bioinformatics. Topics presented include sequence databases, sequence comparisons, database searches, phylogenetic analysis, protein structure, proteomics, RNA structure prediction, gene prediction and identification, genetic analysis of disease, and microarray-based studies of gene expression.

Introduction to Amazon Cloud and UNIX basics (2h)

Description: This workshop is an introduction to Amazon cloud Computing and UNIX basics. The workshop will help you to get setup and build Windows or Unix computer on Amazon cloud computing Attendees will be exposed to hands-on experience and by the end of the course one should be able to confidently use the command line interface on Unix system. They should be able also be to navigate around the Unix file system from the command line and use a number of basic Unix commands.

Analyzing NGS Data: Standard data processing and workflow analysis on high powered computing environment (3h)

Description: The NGS technologies have the potential to dramatically accelerate biomedical research by enabling comprehensive analysis of genomes and transcriptomes to become inexpensive, routine, and widespread tools. This workshop will focus on methods for base-calling and variant-calling, for aligning reads to reference sequences (e.g. genomes), and for de novo assembly of short reads into longer sequences. The following tools will be covered on Orchestra, a shared research cluster; quality reports of FASTQ files, trimming and filtering of reads, alignment and coverage objects such as SAM/BAM files using bowtie/bwa, calling SNPs with Samtools, and De novo Assembly using Velvet.

Introduction to ChIP-Seq and data analysis using Galaxy (3h)

Description: This workshop is focused on analysis of ChIP-seq data using Galaxy. It is aimed at researchers who are using, or planning to use ChIP-Seq methods as part of their research. The course will focus on hands-on training in using standard data analysis methods including loading data into galaxy, Quality control and manipulation, Bowtie mapping, Peak Calling using MACS, Annotation and functional enrichment of peaks.

Navigating and Using NCBI BLAST & Gene Expression Omnibus (2h)

Description: Learn how to use BLAST as an experimental tool. We will cover the use of filters as BLAST tools and contrast them with the use of PHI-BLAST. We will learn about substitution matrices and how the PSSMrelates to PSI-BLAST. The Gene Expression Omnibus (GEO) is a public repository that archives and freely distributes microarray, next-generation sequencing, and other forms of high-throughput functional genomic data. Learn how to navigate the GEO interface to retrieve data to inform your

office for postdoctoral fellows

25 Shattuck Street | Gordon Hall 006 | Boston, MA 02115 | 617.432.6735 | postdoc_office@hms.harvard.edu | postdoc.hms.harvard.edu

experiments. This workshop will acquaint attendees with NCBI blast tool for comparing nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. Attendees will also learn how to query and download experiments that are MIAME-compliant.

mRNA-seq analysis using JMP Genomics Software (3h)

Description: Next-Generation sequencing is quickly becoming the platform of choice for genomic analysis. JMP Genomics has incorporated a number of functions for working with mRNA-seq data. Many of the functions are similar to those used in traditional expression analysis. Learn how to Map and generate counts, import mRNA-seq data (counts) and carry out mRNA-seq analysis including data filtering, normalization, ANOVA and differential expression.

Affymetrix / Illumina Microarray data analysis using R/Bioconductor (3h)

Description: The course is a general introduction to Microarrays and the use of R/Bioconductor to carry out microarray data analysis. Following introduction the workshop starts with hands-on exercise on how to install R and Bioconductor GUI packages. The course is mainly based on the use of Bioconductor open source packages for analyzing single channel and two channel data sets. Only basic R coding will be introduced since all the analysis are performed using AffylmGUI, a graphical interface to Bioconductor tools, designed for life scientists who are not familiar with R language. Students will learn how to carry out the following; Quality control, Normalization, Filtering, Statistical analysis, and Differential expression.

HTqPCR - high throughput qPCR analysis using R/Bioconductor (1.5h)

Description: Quantitative real-time polymerase chain reaction (qPCR) is routinely used for RNA expression profiling, validation of microarray hybridization data and clinical diagnostic assays. HTqPCR, a package for the R statistical computing environment aids to enable the processing and analysis of qPCR data in high throughput across multiple conditions or replicates, and in spatially-defined formats such ABI TaqMan Low Density Arrays and conventional 96- or 384-well plates. The workshop is designed to help researchers learn how to load data into HTqPCR and carryout quality assessment, normalization, visualization and parametric or non-parametric testing for statistical significance in Ct values between features (e.g. genes, microRNAs).

Geneious Pro Genomic Analysis Software Platform (2h)

Description: Attendees will have an opportunity to learn the following. Full genome sequence assembly Sequence, literature & BLAST searching, Phylogenetics, Primer design/Primer Management, in silico cloning and Gateway cloning, Variant (SNP) calling, RNA-Seq Mapping and Expression Analysis

hms | hsdm

office for postdoctoral fellows

25 Shattuck Street | Gordon Hall 006 | Boston, MA 02115 | 617.432.6735 | postdoc_office@hms.harvard.edu | postdoc.hms.harvard.edu

Making the Most of the UCSC Genome Browser (2h)

Description: The UCSC Genome Browser provides rapid, straightforward access to a vast store of Genome- oriented material. Learn how to quickly locate gene information, gene features, how to download sequence and track information, and how to make use of the Table Browser to retrieve data in bulk. We'll also examine other UCSC tools such as the Gene Sorter and VisiGene.

Ensembl Genome Browser Workshop (2h)

Description: Ensembl provides unified access to genomic information and annotation for more than 50 eukaryotic species. Learn how to find what you need, from splice sites to regulatory regions to SNPs. We'll also explore the BioMart tool to select and export Ensembl data using hands-on exercises.

Graphics and statistical tests using R and JMP software (2h)

Description: In this class attendees will have opportunity to learn basic concepts of R and some useful applications of R in research: the ggplot2 graphics and some common statistics tests. This class also exposes attendees with JMP Statistical Discovery Software from SAS. This software provides the complete spectrum of statistics and graphics a student or researcher may encounter. JMP is visual, interactive and dynamic, with a friendly point and click, drag and drop interface. JMP combines powerful statistics with dynamic graphics, in memory and on the desktop. Its interactive and visual paradigm enables JMP to reveal insights that are impossible to gain from raw tables of numbers or static graphs

Pathway Analysis and Networks of Genomic data using Metacore (2h)

Description: How to choose the right network-building algorithm to test and expand your hypothesis. One aspect of systems biology is to integrate complex interactions of biological systems. GeneGo provides a highly annotated and dense interaction database with over ten different network-building algorithms. Here we demonstrate the strength of these tools in the ability to visualize signaling interaction networks and expand on your hypotheses outside of the realm of your core research areas. This tutorial describes each network building algorithm and modeling workflows including building with canonical pathway interactions, with examples of when to use each. In this session we also highlight how to optimize the visualization of your interactions of interest on a network we will build. We will show tools such as how to add/ hide/show objects and how to manipulate visualizations of pathways using post-filters such as disease, tissue, orthologs or gene ontology processes.

Analysis of MicroRNA Expression and Function by a variety of Techniques (2h)

Description: MicroRNAs (miRNAs) are post-transcriptional regulators that silence gene expression by binding mainly to untranslated regions in the 3' end of the target messenger RNA transcripts. This session will be useful for any researcher who wants to investigate miRNA's in detail. Topics covered include basic techniques for miRNA isolation, expression profiling and validation, as well as their functional analysis in mammalian cells.