

Summer 2020

Unix/Linux for Informatic Analysis

https://github.com/sabrsyed/InformaticsTools_2020

Week 3

- Unix Problem Set Q&A
- Review ChIP-Seq and RNA-Seq pipelines
- Loading files to the cluster using FileZilla and rsync
- On the ghpc06 cluster, read criteria for running bowtie2, picard, samtools, macs2, etc.
- See sample bsub commands for running each module
- Look at bowtie2 and macs2 output files
- Use IGV to view .bam and .bed files
- On the ghpc06 cluster, read criteria for running rsem-calculate-expression on the cluster
<https://bioinfo.umassmed.edu/index.php?p=33>
- RNA-Seq – Look at RSEM output files
- See sample bsub commands for rsem-calculate-expression

Tasks:

Create lab schedule for using the ghpc06 cluster

Explore the cluster and examine modules and tools for data analysis that are available

Upload Fastq files to ghpc06 cluster

Week 4

- Introducing NGSplot: Quick mining and visualization of NGS data by integrating genomic databases
<https://github.com/shenlab-sinai/ngsplot>
- Introducing HOMER: Software for motif discovery and next-gen sequencing analysis
<http://homer.ucsd.edu/homer/ngs/>
- Running NGSplot on the ghpc06 cluster for making ChIP-Seq and RNA-Seq heatmaps and extracting tag density files
- Running HOMER on the ghpc06 cluster for annotating bed files and performing motif analysis
- Running bedtools on the ghpc06 cluster for comparing .bed files of different experiments