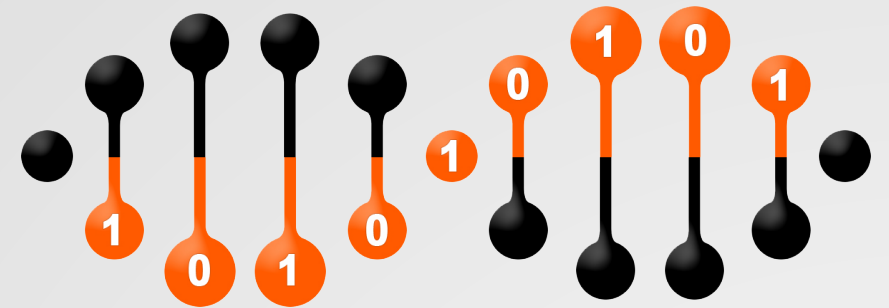


Welcome!

Bi-weekly Data Therapy

Vijender Singh
Neranan Perera
Jill Wegrzyn



Computational Biology Core

UConn
UNIVERSITY OF CONNECTICUT

CBC: Hardware

Two Clusters (Free Accounts)

BBC – Research and Teaching

HPC1 – Advanced Research

Xanadu – New Cluster Coming Online Soon!

- Access to more high memory machines

Hardware

To request an account on the cluster, please [contact us](#). If you are planning a project with the CBC (or the [CGI](#)) and need to include a description of the computational resources, please use this [text](#).

[BBC Cluster \(bbcsrv3.biotech.uconn.edu\)](#)

32-node Dell Linux cluster running Centos 6.3 and Rocks 6.1 x64.

Node configuration:

- 1: 16 x Quad-core 2.5 Ghz AMD 6370P processors with 512 GB RAM
- 4: 4 x 8-core 2.00 GHz Intel Xeon processors with 64 GB RAM
- 17: 2 x Quad-core 2.53 GHz Intel Xeon processors with 32 GB RAM
- 10: 2 x Dual-core 2.8 GHz AMD processors with 16 GB RAM

Local storage capacity: Dell PowerVault MD1000 18TB RAID Array

[UCHC HPC Cluster \(hpc1-submit-int.cam.uchc.edu\)](#)

Node Configuration:

- 2,400 CPU cores with 5.8 TB RAM
- 7,000 GPU cores with large CPU-only and hybrid compute clusters + OSG cluster

Virtualization Infrastructure:

- 456 CPU cores, 1 TB RAM VMWare server and desktop virtualization hosts hosting 100+ Windows/Linux virtual machines with SSD high OPS performance cache tier

Datacenter Infrastructure:

- UPS generator backed power with redundant cooling
- 3x40 Gbe dark fiber connection to off-site DR location

Network (100+ Gbe):

- Full non-oversubscribed 10/40 GbE datacenter network core layer
- BioScienceCT Research Network – 100 GbE to CEN, Internet2, Storrs
- New HPC Science DMZ – low latency, 80 Gb-capable firewall

Storage (2+ PB):

- 394 TB EMC² Isilon and 832 TB Gumulo shared scale-out clusters along with 1,056 TB Amplistor on-premise cloud storage



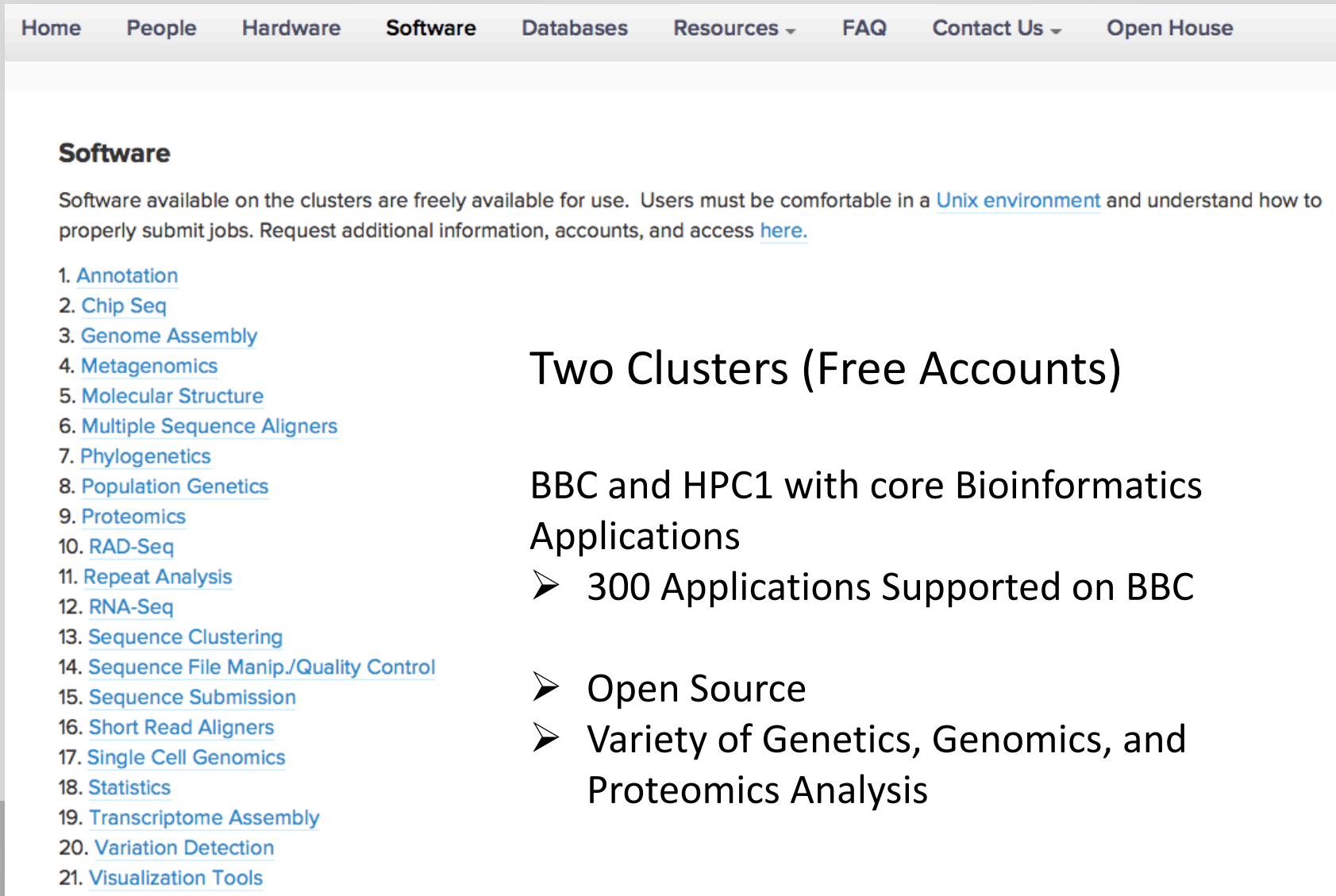
Not our cluster (CERN) but here for inspiration



Institute for Systems Genomics:
Computational Biology Core

bioinformatics.uconn.edu

CBC: Software



The screenshot shows the 'Software' page of the CBC website. At the top is a navigation menu with links for Home, People, Hardware, Software, Databases, Resources, FAQ, Contact Us, and Open House. The main content area is titled 'Software' and contains a paragraph stating that software is freely available but requires a Unix environment. Below this is a numbered list of 21 software categories, each with a blue underlined link. To the right of the screenshot, there is a text overlay with a title and three bullet points.

Home People Hardware Software Databases Resources FAQ Contact Us Open House

Software

Software available on the clusters are freely available for use. Users must be comfortable in a [Unix environment](#) and understand how to properly submit jobs. Request additional information, accounts, and access [here](#).

1. [Annotation](#)
2. [Chip Seq](#)
3. [Genome Assembly](#)
4. [Metagenomics](#)
5. [Molecular Structure](#)
6. [Multiple Sequence Aligners](#)
7. [Phylogenetics](#)
8. [Population Genetics](#)
9. [Proteomics](#)
10. [RAD-Seq](#)
11. [Repeat Analysis](#)
12. [RNA-Seq](#)
13. [Sequence Clustering](#)
14. [Sequence File Manip./Quality Control](#)
15. [Sequence Submission](#)
16. [Short Read Aligners](#)
17. [Single Cell Genomics](#)
18. [Statistics](#)
19. [Transcriptome Assembly](#)
20. [Variation Detection](#)
21. [Visualization Tools](#)

Two Clusters (Free Accounts)

BBC and HPC1 with core Bioinformatics Applications

- 300 Applications Supported on BBC
- Open Source
- Variety of Genetics, Genomics, and Proteomics Analysis



CBC: Tutorials

Custom Guides:

Interacting with the Cluster
Analysis and Viz with R!

RNA-Seq

Genome Assembly

More on the way!

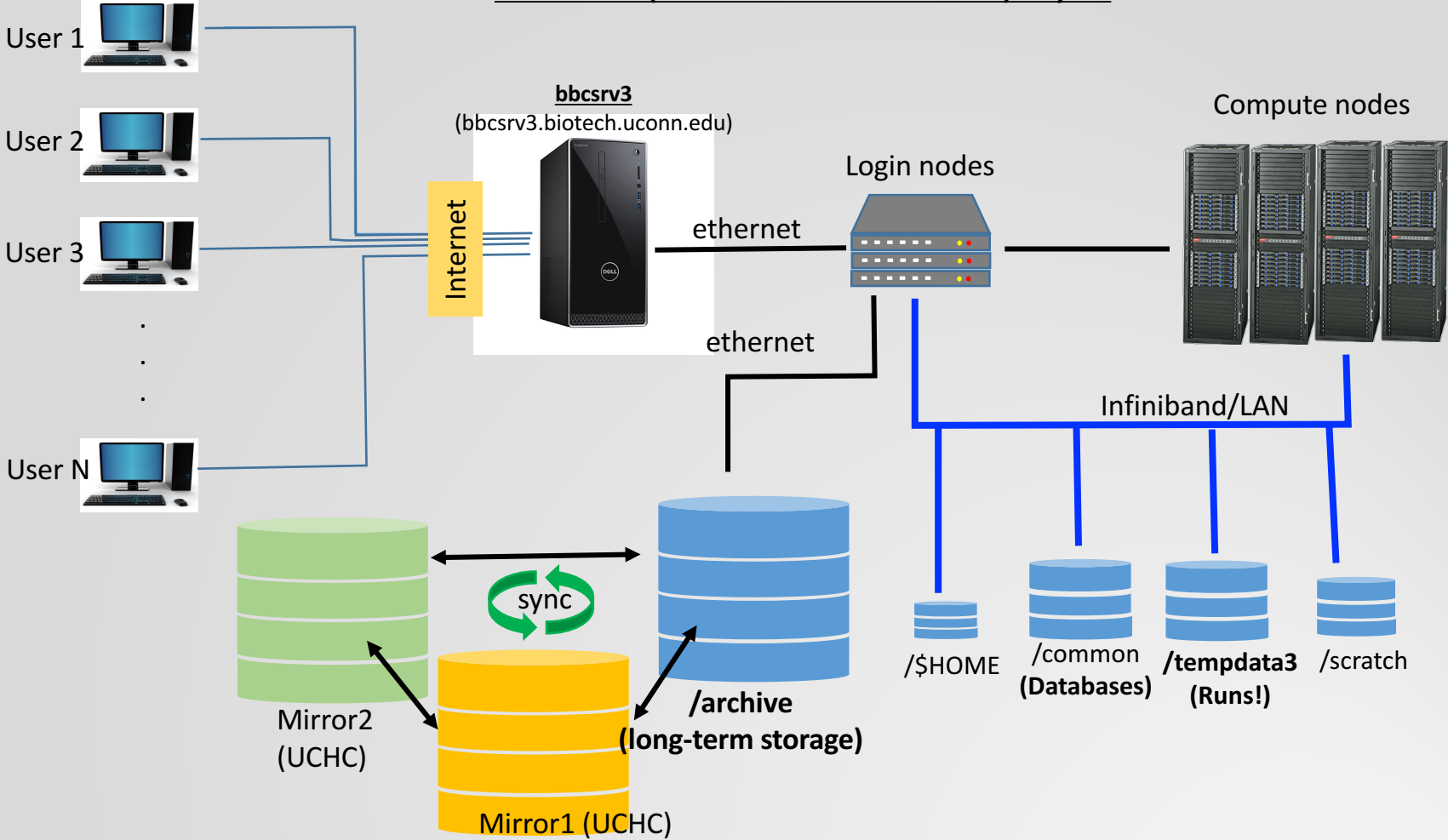
Tutorials

Tutorial	Last Updated	Description
Server Access		
UConn Health Cluster (PBS)	June 2015	Understanding the UConn Health Cluster
BBC Cluster (SGE)	June 2015	Understanding the BBC Cluster
UNIX and R		
Unix Basics	November 2013	Introduction to Command Line Operations
VIM	December 2013	Introduction to VIM UNIX Editor
Unix Examples	November 2013	Basic Bioinformatic Exercises
Introduction to R	November 2013	Basic Analysis and Plots
RNA-Seq Guides		
Prokaryote RNA-Seq (EDGE-pro/DESeq2)	July 2015	EDGE-pro tutorial (with Listeria reference genome)
Model Plant RNA-Seq (STAR/DESeq2)	July 2015	RNASeq tutorial (with Glycine max reference genome)
Non-Model Plant RNA-Seq (Bowtie2/eXpress/DESeq2)	August 2015	RNA-Seq tutorial (with Picea rubens reference transcriptome)
Human RNA-Seq (no replicates) (STAR/DESeq2)	March 2015	Introduction to RNASeq
Model Insect RNASeq (Web-based Galaxy)	July 2015	RNASeq tutorial (with Drosophila reference genome)
Genome Assembly		
Genome Size Estimation Tutorial	January 2017	Genome Size Estimation Tutorial
Bacterial Genome Assembly Tutorial	September 2015	Genome Assembly tutorial



CBC: BBC Cluster

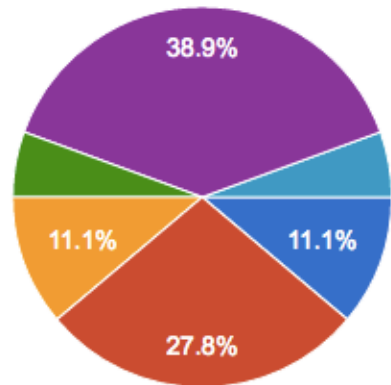
BBC Cluster (bbcsrv3.biotech.uconn.edu) Layout



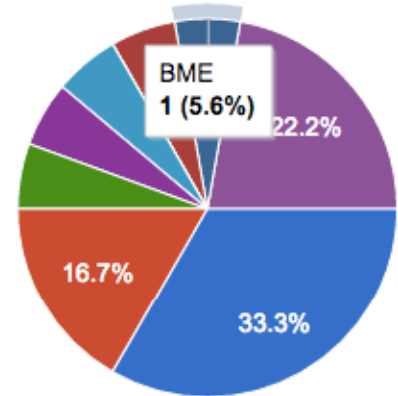
Submitting Jobs!
Where to Run and Store?
Long-term storage (/archive)



CBC: Audience!

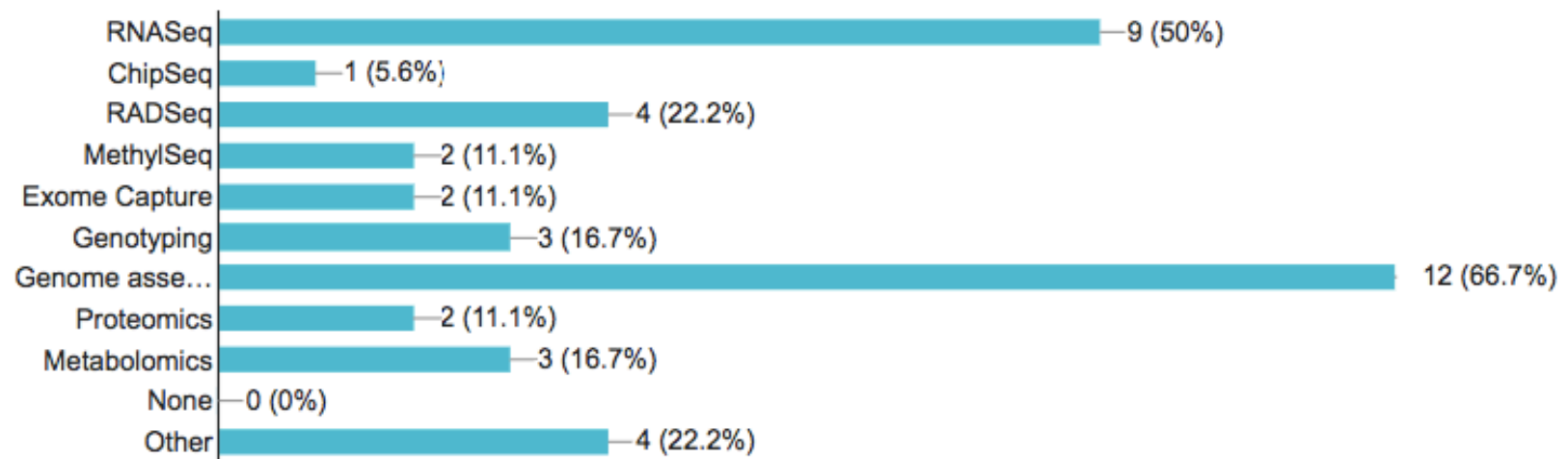
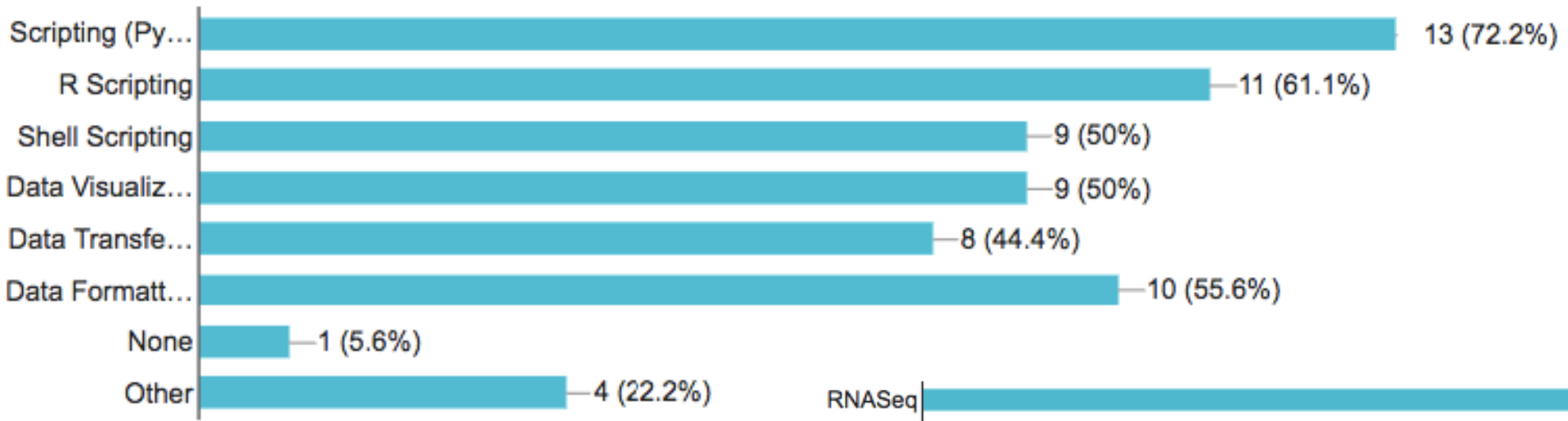


- Undergraduate
- Graduate Student
- Postdoc
- Staff Researcher
- Faculty
- Other



- MCB
- EEB
- PNB
- NRE
- Plant Science
- Animal Science
- Marine Science
- Stats

▲ 1/2 ▼



CBC: Organize

Invited Speakers

Research Talks

Directed Q&A

Code Exercises

Mini Tutorials



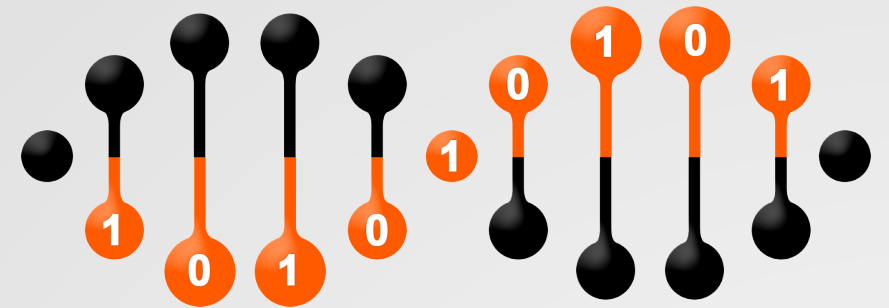
Stay Connected!



@Uconn_Bioinfo



Bioinformatics-I@Listserv.uconn.edu



Computational Biology Core

UCONN
UNIVERSITY OF CONNECTICUT