Bioinformatics Applications and Galaxy

Mark Howison
Application Scientist, CCV

CCV/EPSCoR Bioinformatics Workshop, October 18, 2013
Software modules on Oscar

$ module avail :bio
$ module avail blast
$ module help blast
$ module load blast/2.2.28+

Need to request a software installation?
  support@ccv.brown.edu
Interactive vs. batch jobs

$ interact

  Forwards your session to a compute node
  If your connection is interrupted, the job terminates
  default: 1 core, 4GB memory, 30 minutes

$ interact -n 8 -m 20g

  8 cores, 20GB memory

$ sbatch myscript.sh

  Runs the script on compute nodes in the background
  Writes the output to a file
Essential commands on Oscar

$ myq
   List all of your pending/running jobs

$ allq
   Same, but for the entire queue (all users)

$ nodes
   Summary of available nodes, how many are in use

$ myquota
   How much disk space you have left
Example batch jobs

$ sbatch ~/batch_scripts/blast.sh
   1-core, serial job

$ sbatch ~/batch_scripts/bowtie2.sh
   8-core, threaded job

$ sbatch ~/batch_scripts/raxml.sh
   16-core, distributed/MPI job

$ sarray ~/batch_scripts/blast_array.sh
   Job array: launches several 1-core, serial jobs
Galaxy

GUI web-based platform for bioinformatics
Available as a public server, but limited resources
CCV is deploying a local instance

Interested in being an early adopter?

mhowison@brown.edu
Other resources

BioMed Wiki:
https://sites.google.com/a/brown.edu/bioinformatics-in-biomed

Genomics Club:
https://sites.google.com/a/brown.edu/genomics-club/