

Computational genomics methods (computer lab) ME:800.723
All sessions 12:45-3:15PM

Monday, April 16: *nix and command-line utilities

Objective: understand the Unix environment, and use a small set of Unix command-line utilities to manipulate sequence alignment files.

Overview of directory structures, changing directories, managing files, looking at file sizes and file permissions, write a simple shell script.

For more information: these tutorials will be very helpful if you need more practice (the first one is basic, and quite short):

<https://www.codecademy.com/learn/learn-the-command-line>

<http://swcarpentry.github.io/shell-novice/>

http://korflab.ucdavis.edu/unix_and_perl/ (concentrate only on the unix shell tutorial)

Thursday, April 19: Python

Objective: understand the range of toolsets available in Python

Create simple parsing scripts in Python, manage file input and output, and load modules for more specialized work.

For more information:

<https://docs.python.org/3/tutorial/index.html> (official tutorial, very detailed)

<https://www.codecademy.com/learn/python>

Monday, April 23: running command-line programs, shell scripting

Objective: manage a bioinformatics environment

Run common command-line programs for sequence alignment and peak finding, using shell scripts to manage input and output. Write python scripts for parsing output.

Homework (not graded): swirl

<http://swirlstats.com/students.html>

Thursday, April 26: parsing and processing with Python and R

Objective: write short Python scripts to implement key algorithms and parse files, and use the R statistical programming framework to explore the results

Monday, April 30: R and GRanges

Objective: use R to explore sequence alignment data

Make summary graphs and tables, and use the GRanges package to explore overlaps of the demo datasets with common genome annotations

Thursday, May 3: Data visualization

Objective: understand multiple ways to look at high-dimensional biological data

Create IGV, UCSC, Circos plots and more.

Homework (graded): Use the skills you've learned to tackle several computational genomics challenges. Due May 14.