Biol 381 – Bioinformatics Lab
Lab 1: Introduction (to Python)

Introduction

Why program?

Many steps in the research process depend on algorithms - searching biological literature, predicting sequence homology, identifying gene regions in raw sequence data, building networks of gene interactions . . . Many programs are freely available and accessible online. But what if your research problem requires a different order of steps or many more sequences than can be efficiently entered into a web interface manually or utilizes input files larger than the web service allows? With a few basic programming skills you can write programs specific to your research question and automate the gathering and processing of many, many records at once. Many common programing languages have collections of code written specifically for common biological applications (e.g. Biopython, BioPerl, BioRuby, BioC++, BioJava).

Why Python?

Python is a user-friendly, general purpose programing language. Its freely available, easy to install, and very well documented - and lots of biologists use it! In this course we'll use Ananconda, a free distribution of Python that comes with many useful scientific computing libraries (including Biopython). Anaconda is already installed on the computers in the Gold lab, but if you want to install Anaconda on your own computer you can download it from https://store.continuum.io/cshop/anaconda/.

Part 1: Getting to know Python [group exercise]

As a class, we will work through a series of exercises designed to introduce you to the basic operations, data types and data structures of Python. At the end of class you should be able to assign values to variables; test equality; perform basic math operations; differentiate between strings, lists and dictionaries; test what type of data is stored in a variable; and write basic if statements and for-loops.

Step 1:

Open the iPython notebook Python_intro.ipynb from Blackboard. Rename the file.

Step 2:

Make modifications to the code and add comments to the notebook as we go through the exercise as a class.

Due before the next lab period:

An iPython notebook containing your answers to the questions and commented code.
You will submit this and all other homework assignments via Blackboard. (Don't worry we'll go over this in class.)