

## Lab 4: Intro to Python

### Background

This lab will be the first Python programming lab. Therefore, the “write-up” will be a bit different than previous labs. You will not submit a written report/article. You will only submit your script file(s) (“xyz.py”) for the assignment to Blackboard. No formal report (like the previous labs) is necessary. However, a large portion of your grade will be based on the thorough commenting embedded within your script.

**You may collaborate to some degree with classmates, but it is expected that all code written is your own.** This means you can discuss how to write code, what code does what, or the best way to write code to do a certain task with your classmates, but there should be no copying code or comments.

### Learning Objectives

1. Gain experience writing scripts in python.
2. Gain experience with the basic concepts covered in class including functions, variables, if/else, for/while loops, etc.
3. Gain experience opening a file and parsing data.
4. To produce some code that is functional and might become useful down the road in a bioinformatics context.
5. To practice documenting code with embedded comments.

### Tasks

Develop a python script that:

1. opens the provided nucleotide FASTA file (but could work with any other nucleotide FASTA file). You may assume that the
2. calculates the frequency of the four nucleotides in each sequence and across all samples.
3. compares the frequency of transitions to transversions (where transitions or transversions are clear and unambiguous). You should assume that the sequences provided are perfectly aligned. You may not be able to tell if a transition or transversion occurred at some sites—ignore these sites in your counts.
4. writes to an output file the frequencies of the nucleotides in each sequence **and** across all sequences **and** the proportion of transitions to transversions.

Be sure to annotate your script with LOTS of descriptive & informative comments, including, a comment at the top describing what the software does, how to use

it, required input & expected output, comments defining variables, comments defining functions, etc. Your script should run without producing any error or warning messages when run on the provided FASTA file as well as our different FASTA file that we will use for testing.

## **Guidelines**

Submit your Python script files(s) to Blackboard. These should be formatted as .py files, not PDFs, not Word documents, not anything else. You may attach supplementary files (e.g. FASTA files) if you think it is necessary, but it's probably not.