

Lab03
BIOL638/738

Deliverables

Program filename: nucParams.py

40 total points

Submit your script by 11:59pm on Oct 17, 2021.

For this lab, you will submit a single file named nucParams.py. You can work on whatever Python3 editor you would like (e.g. online-python.com or on nano on Trypsin). In this exercise, you'll create a Python3 program that:

- takes in a FASTA file as input (10pts)
- calculates the nucleotide content (A,T, C, Gs percentages) (10pts)
- Outputs the sequence length (10pts)
- Outputs the GC content (10pts)
- *Extra Credit:* amino acid composition (10pts)

For example, for the given FASTA file (sequence.fasta), the program should output:

% of As: 0.30%
% of Ts: 0.32%
% of Cs: 0.18%
% of Gs: 0.20%
GC Content: 0.38%
Sequence Length: 29903bp

Extra Credit Output:

A = 0.04%
T = 0.07%
I = 0.04%
M = 0.01%
L = 0.09%
G = 0.04%
R = 0.06%
S = 0.08%
N = 0.05%
V = 0.05%
P = 0.03%
E = 0.03%
D = 0.03%
_ = 0.08%
F = 0.06%
Y = 0.05%

$W = 0.03\%$

$C = 0.06\%$

$Q = 0.03\%$

$K = 0.04\%$

$H = 0.03\%$