Fill in with a list of commands

1. Write a python script (script that runs in terminal)! Provide the script! (2.5 pts)

Write a for loop that counts the number of sequences in the example2.fasta file.

Import the example2.fasta file.

2. Write a python script (script that runs in terminal)! Provide the script! (2.5 pts)

DNAseq = 'ATTATGCGACGAGCGAGATCGCGATATTACCC'

In the script make these as functions!

a) Count the number of 'AT repeats'

b) Count the number of 'GC repeats'

c) Count the length of each sequence

d) Count the number of EcoR1 sites (G*AATTC)

e) Replaces all T with U and prints to new file called RNA converted

f) Counts AT and GC content

2. Playing with pUC19c plasmid at NCBI

A) Get pUC19c sequence on canvas or our github.

In script: Read in the pUC19c using python. Remember fasta files contain headers in the first line (">"), one option is to start reading at the second line or empty the first line in your list. In your script. On the line that you read in the fasta, comment the line with the pUC19c ACCESSION #.

B) In script: Join the sequence together (make sure the header is not part of the sequence!!)

C) In script: Count the number of SMAI sites (CCCGGG)

D) Counts AT and GC content

E) Prints them in a final list in terminal or output

G) Apply the restriction enzyme SMAI to split the pUC19c DNA in silico.

- https://www.neb.com/products/r0141-smai#Product%20Information

- You do not need to split the cut site in half, just split at the recognition pattern CCCGGG

- Screen shot neb cutter output from pUC19c