

# Lab 06 - Identify if a genetic mutation exists.

---

200pts pts total.

## Problem

---

Some diseases are the result of a single letter genetic mutation. Cystic Fibrosis (CF) is one of these. Usually CF is fatal by age 28. There is a treatment for CF that is very effective. The problem with the treatment is that it is fatal if given to people that do not have the CF genetic mutation.

We have some genetic data (not actually Stephie's) that we are going to search to see if it has a genetic mutation.

The pattern we are looking for is:

```
ACTGGCTGGCCTAGTGA CTTCCAGCTGCACAGCTATCGACCCAGGGCTGGACAGCCCCTGCCTGGC
```

If the pattern is found then the mutation exists.

The genetic data is 32 meg out of 2.3 gig of a full genetic sequence. To download the data use the link in the Announcement - from your browser and Right-Click on Windows or Control-Click on a mac and get the context menu, then pick "Save Link" in the menu.

The link is:

[http://uw-s20-2015.s3.amazonaws.com/GCF\\_000001405.39\\_GRCh38.p41214\\_genomic.s44a-21.fna](http://uw-s20-2015.s3.amazonaws.com/GCF_000001405.39_GRCh38.p41214_genomic.s44a-21.fna)

Our process for searching this is:

1. Using the function to read a file, `xyzzzy.py` read in the data.
2. Convert the data from a list of strings to a single string. The python function `"".join(list)` will do this for us.
3. Clean up the data - in this case it is mixed case and we have an all upper case pattern. So convert it to upper case. `string.upper()` will do this for us.
4. Search the data to see if the pattern is in the data. `string.index(patter)` can be used to do this. It returns the location of a pattern in a string. If not found it throws an error. This means we will need to put it in a try/except block to catch the error.

Turn in your python program and the output location of the pattern if it is found.

# Copyright

---

Copyright © University of Wyoming, 2021.