Using gc.py

This script takes a fasta file as input, and outputs a tab-delimited table containing two columns: the fasta header and the sequence's GC%, rounded to two decimal places. In creating it, I modified some code from **https://bioinformatics.stackexchange.com/a/2658** written by user Chris_Rands that parses fasta files. I also used the biopython library, which includes various functions for manipulating genetic sequences.
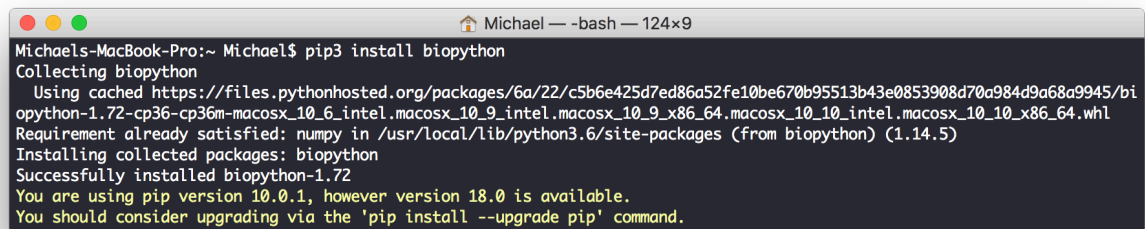
Dependancies

In order to use gc.py, you need to have an installation of python 3, as well as the sys and biopython libraries. The sys library is installed by default. In order to install biopython, use

```
$ pip install biopython
```

> 🔐 If installation is unsuccessful, try using sudo to install your package as root.

You should see something like this:

```
                                🏠 Michael — -bash — 124×9
Michaels-MacBook-Pro:~ Michael$ pip3 install biopython
Collecting biopython
  Using cached https://files.pythonhosted.org/packages/6a/22/c5b6e425d7ed86a52fe10be670b95513b43e0853908d70a984d9a68a9945/bi
opython-1.72-cp36-cp36m-macosx_10_6_intel.macosx_10_9_intel.macosx_10_9_x86_64.macosx_10_10_intel.macosx_10_10_x86_64.whl
Requirement already satisfied: numpy in /usr/local/lib/python3.6/site-packages (from biopython) (1.14.5)
Installing collected packages: biopython
Successfully installed biopython-1.72
You are using pip version 10.0.1, however version 18.0 is available.
You should consider upgrading via the 'pip install --upgrade pip' command.
```

Syntax

The syntax is pretty simple. From the command line, run the python script including your fasta file as the first argument. You redirect the output into the file of your choice. It does not need to be a txt, it can have any extension you would like.

```
$ python gc.py my_fasta_file.fa > output_file.txt
```

> ⚠️ Make sure you are using Python 3. If you are using Python 2, then the program will not work!

Example of Usage

Here is an example of it running. To test this out, I downloaded an example of a fasta file from **http://prodata .swmed.edu/promals/info/fasta_format_file_example.htm** and saved it as `fasta.fa` on my desktop.

```
Michaels-MacBook-Pro:~ Michael$ cd ~/Desktop
Michaels-MacBook-Pro:Desktop Michael$ python3 gc.py fasta.fa > outputfile.txt
```

The output of the file is saved in outputfile.txt on my desktop.

```
outputfile.txt

seq0    6.45
seq1    5.66
seq2    7.46
seq3    10.34
seq4    12.9
seq5    4.55
seq6    11.43
seq7    9.23
seq8    8.82
seq9    8.77
seq10   9.84
```