CSC 314, Bioinformatics Lab #2: **Name:**_____
DNA and Complementary Base Pairing

DNA is a double stranded molecule composed of complementary base pairs. If the sequence of one strand is known, the *complementary sequence* (***complement***) can be determined based on the following rules: adenine (A) binds with thymine (T) and vice-versa; and cytosine (C) binds with guanine (G) and vice-versa.

If a sequence is read from its 5' to 3' end, its ***reverse*** is the same sequence read from its 3' to 5' end (and vice-versa).

The ***reverse-complement*** of a DNA sequence is the reverse of its complement.

1. Find the specified sequences based on the sequence below. Note that your answers must include labels for the 5' and 3' ends. This is the **original** sequence:

   5'-TAGAC-3'

a) Find the reverse of the original sequence

b) Find the complement of the original sequence

c) Find the reverse-complement of the original sequence

d) If the original sequence (5'-TAGAC-3') was an RNA sequence instead of a DNA sequence, what would its sequence be?

2. Complete the Lab 2 Python Jupyter Notebook that prompts the user to enter a sequence, formats the sequence for invalid characters, outputs the sequence and is length, and finds the complement, reverse-complement, and GC content of the specified sequence.

3. Respond to the questions posted on Piazza, related to the Implications of Cheap Genomic Sequencing.