

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. Consider the following DNA sequence. *Make sure to label the 5' and 3' end in your answers to the following questions.*

5'-ACGCCCTGCGCATAAGTG-3'

- a) Find the reverse sequence
  - b) Find the complement
  - c) Find the reverse-complement
  - d) If the original sequence above was an RNA sequence and not a DNA sequence, what would its sequence be?
2. Write a Python program that prompts the user to enter a DNA sequence, from its 5' to 3' end, and then outputs the following:
    - a. The length of the sequence
    - b. The sequence entered by the user (formatted for invalid characters – see below).
    - c. The complementary sequence (don't forget to label the 5' and 3' ends)
    - d. The reverse complement (don't forget to label the 5' and 3' ends)

Note #1: Your program should work for both lower and upper case letters. This perhaps is best accomplished by converting the user's sequence to uppercase.

Note #2: For (b) - (d), if a character in the sequence is not valid, then the invalid character should be replaced with a '-'

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