<div align="center">

Molecular Biology Through Discovery
Problem Set 4: DNA Structure and Replication
Submit reasoning and process, not merely answers!

</div>

1. You work at the Center for Disease Control and have been given to analyze a suspected bioterror agent that may be a type of Variola, a double-stranded DNA virus that is the causative agent of smallpox. Smallpox was one of humanities major scourges, killing as many as 500 million people in the 20th century. It is believed to have been wiped out in 1979 with an aggressive vaccination campaign. Since few alive today are resistant to smallpox, a live virus could wreak havoc.

   To analyze the virus, you obtain DNA from the precious sample of purified virus and (as a control) from canine parvovirus, a single-stranded DNA virus you happen to have on hand. Unfortunately, you forget to label the two tubes! What else to do but to continue,… so you analyze the two DNA samples for nucleotide composition (in the same way as Chargaff), giving the results shown below. Which tube contains which Variola DNA, and why?
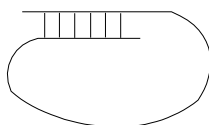
	G	A	T	C	Recovery
Tube 1	15.7	33.4	32.7	11.4	93.2
Tube 2	18.7	34.4	26.9	14.1	94.1

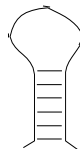2. Consider the sequence below of a single-stranded DNA fragment.

   5'-AGAGAGAGCTAAGGTCTCTCC-3'

   Which of the following is a likely structure for the single-stranded fragment to assume?
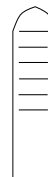

3. You isolate a new restriction enzyme EcoRX (an enzyme that cuts DNA at a specific position). You find that the enzyme cuts the sequence ATGGTATACTGAACGAA once. Like many proteins that bind to DNA, restriction enzymes generally recognize palindromic sequences, usually 4 to 8 nucleotides in length. Presuming that EcoRX does so as well, what is the sequence it recognizes and cuts?

4. The haploid human genome has about 3 billion base pairs, divided amongst 23 chromosomes. Take an average human chromosome and model its DNA as a piece of uncooked spaghetti, with a width of 1 mm and a length you will calculate. How long would it be? If one end were at the entrance to the VCU library and the other end were due north, where would that end lie?

5. A linear piece of DNA has just finished replicating from an origin of replication approximately in the center of the molecule. Draw a simple picture of the finished replication products in a mutant *E. coli* strain that lacks the enzyme DNA ligase. Provide zoomed in views of regions of interest to illustrate pertinent detail. Also provide a legend to your drawing and be sure to indicate the directionality of each strand. (If you have to make any assumptions to answer this question, please state them.)

6. You have an idea of how to accomplish something that nature has not been able to do: make an enzyme that can convert atmospheric N_2 to NH_3 in the presence of oxygen. This would have enormous practical implications. For example, it might be possible to engineer the protein into crop plants and eliminate the need for fertilizer!

The first step is to get a copy of a gene that encodes the normal, oxygen-sensitive enzyme (called nitrogenase). The cyanobacterium Anabaena PCC 7120 (nicknamed A7120) has such a gene, called ALL1440. Take a look at the sequence of the genome near the gene by going to CyanoBIKE and using the SEQUENCE-OF function to display the sequence of A7120. Then type ALL1440 into the **Go to** box and press **Enter**.

Your plan is to use Polymerase Chain Reaction (PCR) to amplify the gene in sufficient amounts to clone it… Don't know how PCR works? Try the following sites (you can easily find many more resources on your own):

basic description: https://www.promega.com/resources/multimedia/pcr/introduction-to-pcr/
more flash: http://www.youtube.com/watch?v=2KoLnIwoZKU

6a. From the genome sequence you displayed, devise primers that will enable you to use Polymerase Chain Reaction (PCR) to amplify the gene, *from its start codon to its stop codon, and nothing else*. To check to see if your primers work, use the RUN-PCR tool, available in CyanoBIKE by bringing down the RUN-FILE function from the INPUT-OUTPUT menu, type in the entry box "run-pcr.bike" (press **Enter**), select the SHARED option, and execute the function.
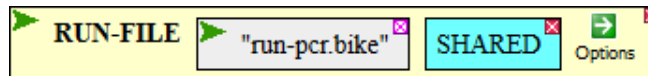
This will bring the RUN-PCR function into your FUNCTION menu. Bring down the function from the FUNCTION menu, enter the sequences of your two primers (between "…"), and enter A7120 as the genome. Then execute the function. If your primers are good, you should amplify the gene ALL1440 and nothing else. *Provide the sequences of the primers and the output of the function.*

6b. Modify your primers so that you get about 10 amplification products, including ALL1440. *Provide the sequences of the primers and the output of the function.*