CSC 314, Exam I Review

Note: the format of Exam I will be similar to the format of previous Labs, but may also include multiple choice, short-answer, or matching questions. The material will cover biological terms and concepts including inheritance, cells and DNA, nucleotide sequences, protein sequences, gene expression, and python programming.

Biological concepts

- *Inheritance*
    - Law of Segregation
    - Law of Independent Assortment
    - Alleles
    - Punnet Squares
    - Homozygous vs. Heterozygous
    - Dominant vs. recessive traits
    - Phenotype vs. Genotype
    - Complexities
        - Incomplete Dominance
        - Codominance
        - Multiple Alleles
        - Pleitropy
        - Epistasis
        - Polygenic Inheritance
    - Chromosomal basis of inheritance:
        - Homologous Chromosomes
        - Linked Genes
        - Sex Determination and Sex-Linked Genes
        - Recombination and crossing over
- *Cells and Cell Structure*
    - Differences between eukaryotic and prokaryotic cells
    - Function of: plasma membrane, nucleus, mitochondria, ribosomes, Golgi apparatus, endoplasmic reticulum
- *Overview of Biological molecules*
    - Polymer vs. Monomer, DNA and protein
    - Dehydration synthesis vs. hydrolysis
- *Nucleic Acids*
    - DNA vs. RNA, 5' vs. 3'
    - Nucleotide components
    - Complementary base pairing
- *Proteins and protein structure* - primary, secondary, tertiary, and quartenary structure

- *Gene expression*
  - Transcription
    - primary transcript and pre-mRNA
    - sense strand vs. anti-sense/template strand
    - promoter, TATA box, RNA polymerase, direction of transcription
    - RNA processing, introns/exons, 5' cap, poly-A tail
    - Alternative splicing
  - The Genetic Code (including using the table to translate a DNA/mRNA sequence)
  - Translation
    - tRNA and anti-codons
    - ~~ribosomes: A, P, and E site~~
    - termination and release factor
    - polyribosome
    - signal sequence
  - Mutations
    - Substitutions: silent, missense, and nonsense mutations
    - Insertions and deletions, frameshift mutations

Python programming concepts

- Printing text or variable values
- User input (*input* function)
- Arithmetic
- Slicing and iterating through strings and lists
- String methods/operations – *upper, count, find, in, replace, len, strip, split*
- *if, elif, else* statements
- defining and using functions
- for loops and *range* function
- reading text from files (*open* function, *for* loops, reading methods)
- dictionaries
- lists and string immutability

1. Assume that for a particular type of dog, brown fur is dominant to white fur, and the gene is *not* on a sex chromosome. A male who is homozygous dominant mates with a female who is heterozygous.
    a. What is the probability that a male puppy produced from this union has brown fur?
    b. What is the probability that a female puppy produced from this union has brown fur?

2. Repeat the previous problem, but assume that fur color is a sex-linked trait (the gene is on the X chromosome). The male parent has brown fur and the female is a carrier. Aside: dogs have 39 pairs of chromosomes, but have the same sex chromosomes (X and Y) as humans.

3. If a male with type A blood (genotype $I^A i$) mates with a female having type B blood (genotype $I^B I^B$), what are the possible blood types of the offspring, and what are the expected genetic ratios?

4. Find the complement of the following DNA sequences (make sure to label the 5' and 3' ends)

    a. 5'-GCGGTAAGCA-3'

    b. 3'-GCGATGAGCA-5'

5. Specify the mRNA sequence that would be transcribed from the following DNA *template* strand:

    5'-TCAATGAACGATCAT-3'

6. Assume that above sequence, on the *template* strand, has a mutation from A->G indicated in bold.

    5'-TCAATGA**A**CGATCAT-3'

    Would the resulting mutation be a missense mutation, a silent mutation, or a nonsense mutation? You must justify your answer to receive credit.

7. Write a python script that reads in nucleotide sequences from two files, *sample1.txt* and *sample2.txt*. The script outputs the length of each sequence and then outputs which sequence is longer. Note #1: You may assume that each file contains the sequence and nothing else (i.e., there is no header). However, the sequence may span multiple lines.

8. Write a function in python that takes a nucleotide sequence as input, and returns a list containing the number of A's, C's, G's, and T's. Note that the sequence may be in either lower- or uppercase. Use this function to output the number of A's, C's, G's, and T's in the sequence AGTGAGAGTAG.

9. Suppose that a python dictionary exists and is of the form

```
translate = {"AUG": "Met", "CUU": "Leu",
             "UCU": "Ser", "UGU": "Cys",
             "UGA": "Stop", …}
```

That is, a key exists for each codon and its corresponding value is the amino acid determined by the genetic code.

Using the dictionary above, write statements in Python in order to translate an RNA sequence stored in a string named *RNA*. Hint: you will need to use the *range* function in order to iterate through the index values that begin each codon, i.e., 0, 3, 6, …