

## CSC 314, Exam I Review

Note: the format of Exam I will be similar to the format of previous Labs, but will 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.

You will be allowed to bring *one* page of notes (front and back, handwritten or typed) to the exam. A copy of the genetic code table that is available in your notes will also be provided. The use of computers or cell phones will not be allowed.

### Biological concepts

- *Inheritance*
  - Alleles
  - Punnett Squares
  - Law of Segregation
  - Law of Independent Assortment
  - Homozygous vs. Heterozygous
  - Dominant vs. recessive traits
  - Phenotype vs. Genotype
  - Complexities
    - Incomplete Dominance
    - Codominance
    - Multiple Alleles
    - Pleiotropy
    - Epistasis
    - Polygenic Inheritance
  - Chromosomal basis of inheritance:
    - Homologous Chromosomes
    - Linked Genes
    - Sex Determination and Sex-Linked Traits
    - 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, RNA, 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 quaternary structures
- *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 and/or variable values
- User input (*input* function)
- Arithmetic
- Slicing and accessing single elements of sequences (strings, lists, and tuples)
- Iterating through sequences
- String methods/operations – *upper, count, find, in, replace, len, strip, split*
- *if, elif, else* statements
- for loops and *range* function
- dictionaries
- lists and string immutability

***Note: The questions below are not a comprehensive review, but contain exercises based on concepts that may appear on the first exam. In addition to doing these exercises, you should look over previous labs and your notes.***

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 phenotypic and genotypic ratios of blood types for the offspring?
4. Find the complements 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 the same sequence has a mutation from A->G, at the position indicated in bold and underlined.  
5'-TCAATGA**A**CGATCAT-3' (wild type)  
5'-TCAATGA**G**CGATCAT-3' (mutant)

Would the resulting mutation be a missense mutation, a silent mutation, or a

nonsense mutation? Note that the *template* strand is shown. You must justify your answer to receive credit.

7. Write a python script that prompts the user to enter 2 nucleotide sequences. The script outputs the length of each sequence and then outputs which sequence is longer. Note: The user will enter each sequence 1 at a time.
8. Write a python script that prompts the user to enter a nucleotide sequence, and outputs the number of A's, C's, G's, and T's. Note that the sequence may be in either lower- or uppercase. Output the number of A's, C's, G's, and T's in the sequence AGTGAGAGTAG.
9. **Finding start codons.** Suppose that a DNA sequence, written from 5' to 3', is stored in *dna*, and its complement (from 3' to 5') is stored in *comp*. Write python code that determines whether one, none, or both DNA strands contain a *start* codon (ATG). To do this, you need to search the original strand as well as its reverse complement (since codons are read from 5' to 3'). Output one of the following: *no start codons found, both strands contain at least one start codon, original DNA strand contains at least one start codon, the reverse complement contains at least one start codon*.
10. 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. Write Python code that uses the dictionary to do the following:

- a. Output the amino acid corresponding to the first codon stored in the variable *RNA*.
- b. Translate the entire RNA sequence. If a *stop* codon is found, you should use the statement `break` which will exit the loop.

Note: for (b) you should use the code below, which iterates over index values in multiples of 3. In other words, *i* will take on the values of 0, 3, 6, etc, which corresponds to the index of the first nucleotide for each codon.

```
for i in range(0, len(RNA), 3):  
    ...
```