Fall 2022 CS/BIOL 123A Bioinformatics Project Tasks

Your group consists of three non-programmer types and one programmer type. Although all of you will be working on the same project, below I have divided the tasks between the non-programmer and programmer types.

YOUR PROJECT TASKS INCLUDES:

NON-PROGRAMMER TYPES:

1. Identify the monogenetic disorder that causes Neurofibromatosis I. Reference the SNP/SNV and Clinical Variation information from NCBI.
2. Distinguish Neurofibromatosis I from the remaining two types of Neurofibromatosis.
3. Describe the Mendelian inheritance pattern associated with this disorder.
4. Identify the metabolic pathway(s) that is involved with this disorder.
5. Describe the mechanism/method/process by which the disease starts and progresses.
6. Identify the diagnostic criteria for Such as from the following literature ref.

**NEUROFIBROMATOSIS** TYPE **I** AND ITS DIAGNOSTIC CRITERIA: A CLINICAL OBSERVATION.

Academic Journal (English) ; Abstract available. By: Purdenko TI; Delva MY; Ostrovskaya LI; Tarianyk KA; Sylenko HY; Pushko OO; Purdenko SV, Wiadomosci lekarskie (Warsaw, Poland : 1960) [Wiad Lek], ISSN: 0043-5147, 2022; Vol. 75 (5 pt 2), pp. 1408-1414; Publisher: Aluna Publishing; PMID: 35758466, Database: MEDLINE

1. Identify the NF1 gene variants that lead to Neurofibromatosis I. Identify the genotypes and phenotypes associated with the variants.
2. Use the PharmKB to identify any FDA approved drugs for Neurofibromatosis I. What, if any, are adjustments to drugs that must be made due to genotypic profiles. <- This info is in the PharmKB for some drugs, but not all.

PROGRAMMER TYPES:

1. Modify the Python module, provided to the class for Programming Assignment 1, to download the FASTA and SNP info for the genes identified in step 7 for the NON-PROGRAMMER TYPES.
2. Modify the DNA sequences with the corresponding SNPs for each gene.
3. Use the following Biopython packages to extract metrics about the proteins the genes express.

from Bio.SeqUtils.ProtParam import ProteinAnalysis

from Bio.SeqUtils.ProtParam import ProtParamData

from Bio.Seq import Seq

You will compare the metrics from analyzing the WildType sequences (i.e., the un-mutated sequences) against the metrics for the mutated sequences. Ping me and I will send you example code.

1. If there is sufficient distinction in step 3 above, then build a ML model to predict if Neurofibromatosis I might develop for a given nucleotide sequence with certain types of mutations.

NON-PROGRAMMER TYPES + PROGRAMMER TYPES :

1. Combine and integrate what each type has found into a final project report.