**Post Genomics – Fall 2025**

**Homework 4**

**Due: Thursday, September 25, 2025**

Explore SNPnexus and FATHMM-XF web server and get familiar with their functionalities. Submit example files on the server for practice, downloading the output files to familiarize yourself with the results. For SNPnexus, you get two outputs (pervariant.txt and a zipped annotation folder), focus on CADD, PHRED, SIFT, and pathway.txt. Make sure when submitting the data you select the correct human genome reference (GRCh38 or GRCh37)

**Part 1 SNPnexus** *(50 points)*

Take the merged CSV files you produced in Homework 3.

* 1. Generate the input file for SNPnexus and submit it on their web server. SNPnexus input requires the strand direction (1, forward or -1, reverse), do both for each variant. Currently, there is a batch submission limitation to 10,000 variants, I suggest writing a python script that writes on a text file the exact format needed for the input. If there are more than 10000 variants to submit, then using terminal (Mac), or command prompt (Windows) to split the main text file into new separate smaller text files using:

split -l 10000 (filename)

From this you will need to submit each smaller-input text file onto SNPnexus separately.

* 1. Download results of text files, per Variant and Annotation, and combine columns we need to the CSV files.

**Part 2 FATHMM** *(50 points)*

Take the merged CSV files you produced from Part 1.

2.1 Generate the input file for FATHMM-XF and submit it on their web server.

* 1. Download results and combine columns we need to the OMI file.

**What to Submit:**

1. **The finished CSV files from Part 2. \*\*Submit on Blackboard\*\***
2. **A single Python file with code for Part 1 and 2. \*\*Submit on GitHub\*\***

**\* Name the files (Your Last Name) \_HW4.csv and (Your Last Name)\_HW4.py \***