**GCBA915: Final Project**: Due on 12/14/13, Saturday by 5pm.

Instructions:

1. Similar to the midterm, the Excel spreadsheet emailed to you contains individual assignments to questions from 1-4. Questions 5 and 6 are common to all.
2. Your final project should contain a single Word file. The Word file name MUST start with your name. Please email it to babu.guda@unmc.edu

**Question 1: (5 points)**

List the gene symbols and names (if available) that are present on the chromosome band assigned to you. If more than 10 genes exist, just show any 10 genes (Use the Map Viewer Tool at NCBI)

**Question 2: (10 points)**

Carryout Gene Set Enrichment Analysis (GSEA) using the data sets and gene sets assigned to you. Cut and paste the GSEA report file into your WORD file.

**Question 3: (5 points)**

Calculate the probability of a given sequence matching the HMM model shown in the class. Use the sequence assigned to you for this calculation.

**Question 4: (5 points)**

Using the UniProt ID assigned to you, get the protein sequence and search it against the Pfam database. List all the unique significant Pfam matches and their functions.

Note: List ONLY the significant matches, and list one type of domain only once.

**Question 5: (5 points)**

Consider the following non-synonymous mutations observed in a tumor genome. State how each mutation can potentially impact the structure/function of a protein (without any extraneous information). Give correct reasoning for each mutation based on the physicochemical properties of the amino acids.

1. V 🡪 L
2. C 🡪 W
3. K 🡪 D
4. G 🡪 W
5. P 🡪 S

**Question 6: (10 points)**

Run the entire NextGen sequence analysis pipeline for the paired-end short read sequences and generate the vcf file. Raw sequence files can be downloaded from the server using the command “cp –rf /storage/share/final\_project/ ~”. The rest of the files are as used in the class session.

How to submit: Copy the final vcf file(yourname.vcf)using the following command:

“cp yourname.vcf /storage/share/project/submit/” You can go to this folder and double check if the file is copied or not.