Chaojie (James) Yuan
Advanced Practical Computer Concepts for Bioinformatics
Final Project Proposal

Description:
I have two ideas for my final project.
Idea #1: My first idea is to create a web interface that accepts a gene identifier from the user, and it will search on NCBI Entrez and return important information, such as the DNA sequence, base content, GC content, RNA sequence, protein sequence, and finally, top 10 results for blastn and blastp. There are tools to do one or a few tasks listed above, I wanted to do all of them in my web interface.

Idea #2: My second idea is to create a web interface that takes a query from the user about SNPs in the CFTR gene, and returns associated SNPs according to the query. Specifically, SNPs in the CFTR gene will be stored in a table on a mysql database. The user can search by typing characteristics and attributes associated with SNPs, such as exonic, intronic, synonymous, nonsynonymous, regulatory, etc. Then, the web interface will put all SNPs associated with that attribute into a table up for display.

Methodology:
For Idea #1, I will have an HTML file accepting the gene identifier and giving it to the CGI script to run. I will be using the Biopython package in my CGI script. It has useful functions such as doing a search on the NCBI entrez database, transcribing the DNA sequence and translating the RNA sequence, etc. It is a package that runs quickly. Then, I will have another HTML template file to display the result and a CSS file to make everything more visually neat and aesthetic, compared to the plain text output.

For Idea #2, I will create a new mysql database on the server with SNPs in them. I will probably not be able to store all CFTR SNPs in them, but I will try to store as many as I can. Then I will make a HTML file that first shows what keywords (attributes and characteristics of SNPs) the user can use to query, then accepts the user query and passes it to a CGI script. The CGI script then does a query with that keyword in the mysql database, and returns the results. I will then have another HTML template file to display the results and a CSS file to make everything more visually neat and aesthetic, compared to the plain text output.

Concerns:
General concern: Do I have to use all the technologies and methods listed in the prompt? In my opinion, it would be hard to use mysql for the first idea since I will not be downloading a source data file, instead, I am accessing the NCBI database and running BLAST online.

<u>About using JavaScript and JQuery</u>: I am not sure about how I should incorporate Javascript and/or JQuery in either of my ideas for the project. Dr. Orvis, would you give me some advice on where and when I should use them?