

CSC 314, Methodology Assignment

Cancer Biology Collaboration

Based on the summary and request from your collaborators (the cancer biology students), you will need to come up with a plan (a methodology) for carrying out the requested task.

After discussing the task and possible methodologies with your group, you will need to turn in the following:

1. Clearly explain the objective of what you need to do, and what information you will be providing to your collaborators.
2. Clearly explain how you will accomplish the above task. Specifically, state what libraries, databases, and tools you will need to use, and for what purpose. For example, if you will be downloading sequences, which database or tool will you use to get the sequences from? You may list these steps in "recipe" form, though note that after the experiment you will be writing up your methods and results in paragraph form.

The following is a simple example for a different analysis:

1. The objective of the analysis is to get a list of all human HBB sequence entries from GenBank and their lengths.
2. Methods
 - a. The Entrez package of Biopython will be used to query GenBank (the 'nucleotide' database searching for 'HBB[Gene Name] and "Homo sapiens"[organism]. In this step, the *Entrez.esearch* function from the *Entrez* Biopython module will be used to get a list of records, and to find the total number.
 - b. We will then use the *Entrez.efetch* method to retrieve each sequence in FASTA format, and find the length of each sequence. The ID and length of each sequence will be written to a file.