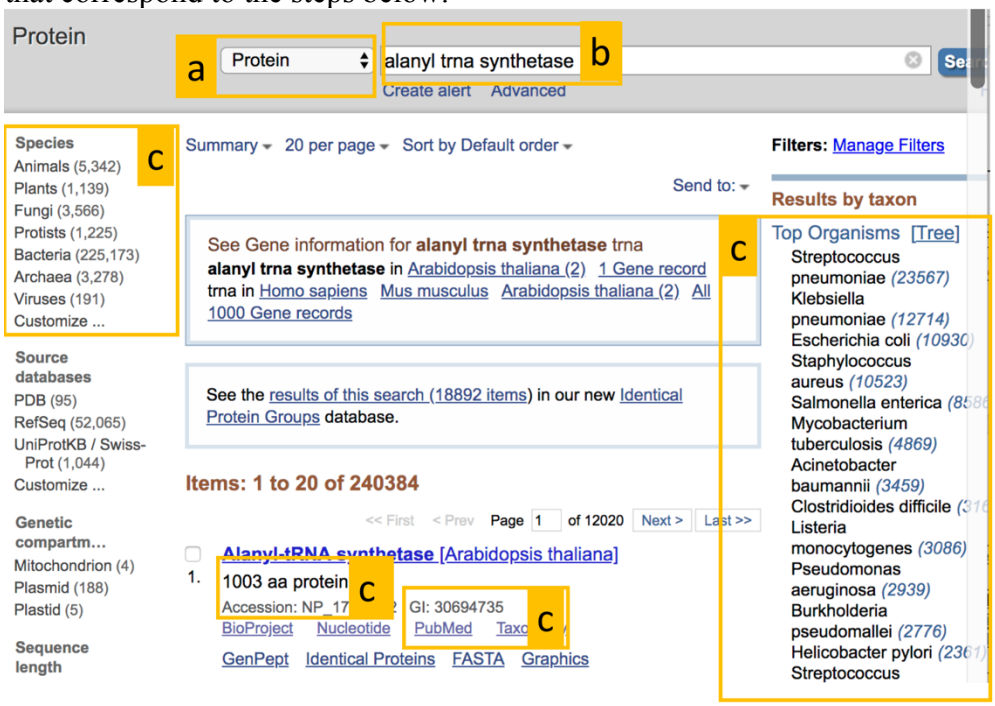
Use the NCBI data base to build a phylogenic tree of at least 5 organisms using proteins sequence. I will describe below how to do this with proteins, but you are free to try nucleotides as well, for example, ribosomal RNA or tRNA. If you are competent with computer programming, there are many other mechanisms to do this. **The end result should be two phylogenic trees each representing the same 5 (or more) organisms: 1 based on a single protein/gene sequence, the other based on organismal level phylogenies.** See the example below.

The guide below is to help generate a tree using the simplest method possible. Follow along with the included image for further guidance. There are some recommended proteins and organisms to get you started, but feel free to deviate from the list.

Proteins to use	Organisms to use
Cytochrome C	Arabidopsis
ATP synthase	Drosophila melanogaster
Ribosomal protein L5	Escherichia coli
Alanyl tRNA synthetase	Pseudomonas aeruginosa
cytochrome c oxidase	Shigella sonnei
	Pyrococcus furiosus
	Homo sapiens

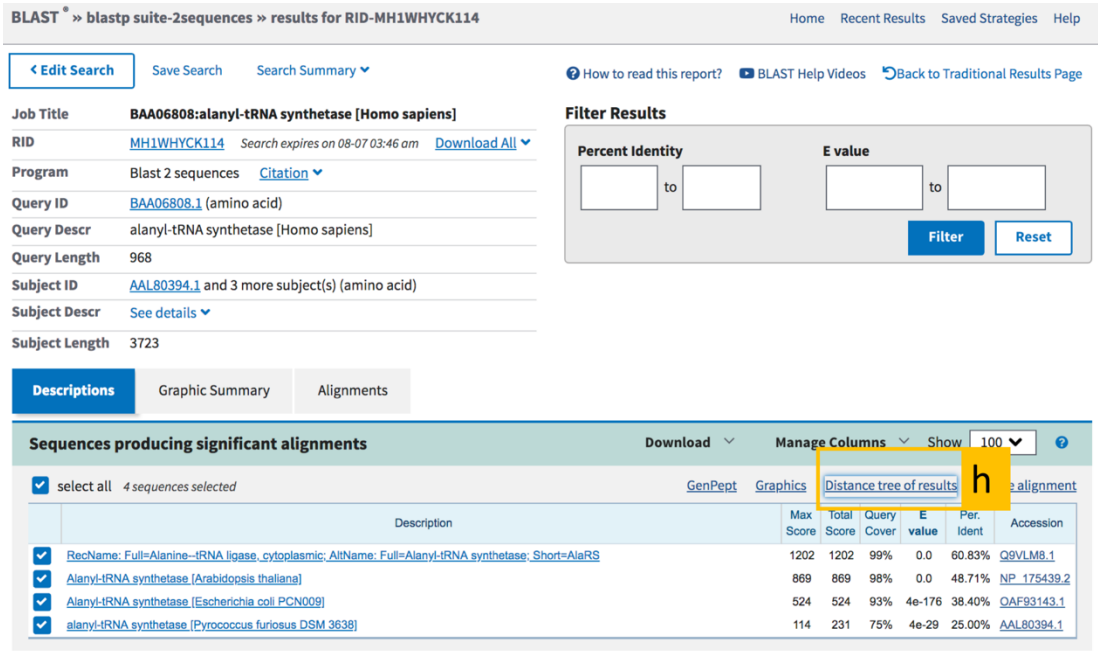
Generate a phylogenic tree from a single protein/gene. Website images are labeled with letters that correspond to the steps below.



a. Go to https://www.ncbi.nlm.nih.gov/ and change the search dropdown menu to "Protein"
b. Search for a protein that you wish to generate a tree for (either from the list, personal interest, or something mentioned in the lectures)

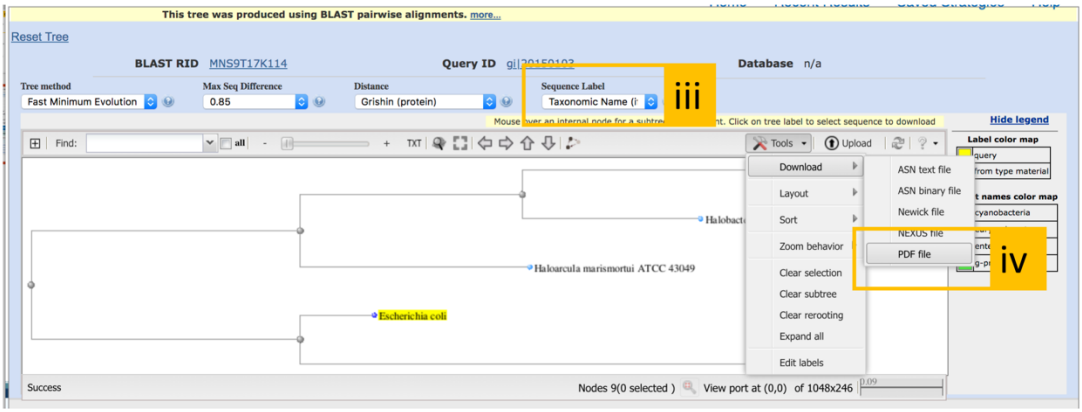
i. When choosing a sequence, some understanding of the protein chosen will help generate a better tree. Note that proteins found in deep evolutionary history are more likely to be characterized for most organisms. For example, if you pick a protein like hemoglobin, it is unlikely to be found in bacteria; ATP synthase is likely to be found across phylogenic clades.

c. Record the GI number of the protein in different organisms, such as the organisms on the list. Try to make all of the lengths (indicated by aa) close to one another (within 20 amino acids). Once you have 5 or more GI numbers go to the next step.
   i. Note that some will be "PREDICTED." This means they sequenced a genome and think this will be the resulting protein. It is fine to use these, but it may not fit the expected result as well.
   ii. Some proteins have "isoforms" meaning there are several types of that protein in the organism. Use aa length to choose one.
   iii. Some proteins will have "partial" or "subunit" these are unlikely to match a full protein, so avoid these sequences
   iv. In the top left you can choose bacteria, archaea, animals, etc. It may be interesting to pick a few from different groups.
d. Go to https://blast.ncbi.nlm.nih.gov/Blast.cgi and select protein blast.
e. Put **one** GI number in the first box, then click on the "align two or more sequences" option. It doesn't matter which one for your tree. It will impact some of the other data you will be given, that you do not need to use.
f. Put the remaining GI numbers in the second box. Click "BLAST."
g. The results page will appear after processing. With only 5 sequences this should be pretty fast. On the results page you will have a table that compares sequence identity of the GI from the first box with the other sequences entered.
h. Generate a "distance tree of results"

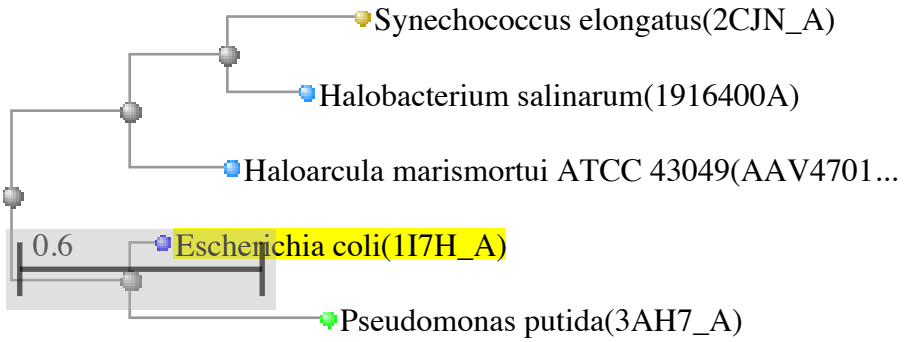


i. In this window, you can see your tree based on the protein you used.
   i. The "tools" dropdown menu will allow you to change the type of tree.
   ii. You can expand or collapse nodes by clicking on them

iii. The fourth dropdown menu defaults to “sequence title”, change it to “Taxonomic name”
iv. To export in the form shown below, go to tools>download>pdf file)



Example of a tree from ferredoxin sequences:



Next generate a phylogenic tree based on taxonomy. Using the organism names (e.g. Homo sapiens) from your protein tree, create a taxonomic tree. I used a free website to do this: https://phylot.biobyte.de/. (It did not like some of the copied names, and I needed to type them in individually)

Example tree from phylogenic information: This is just a screen shot of the tree.