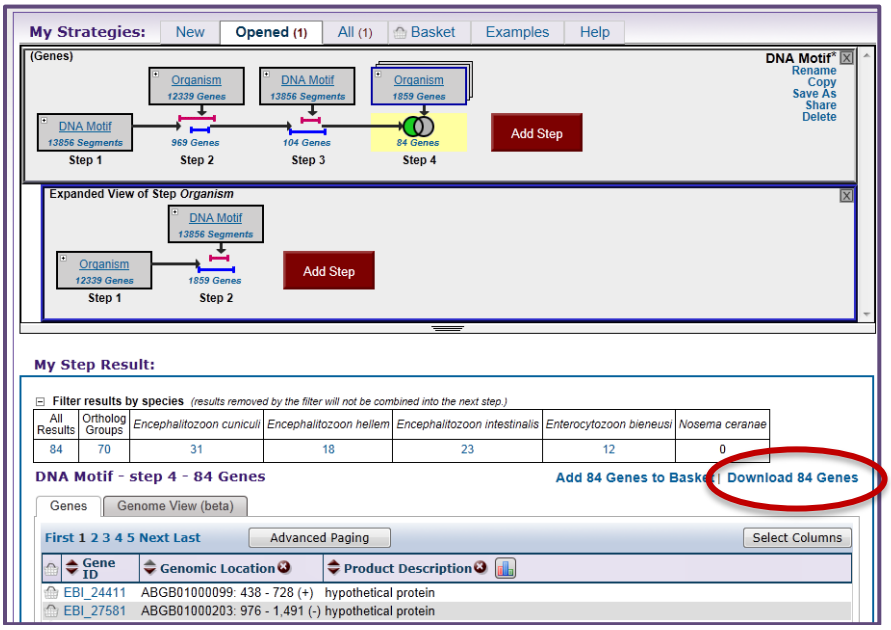
Data retrieval and download
Exercise 9

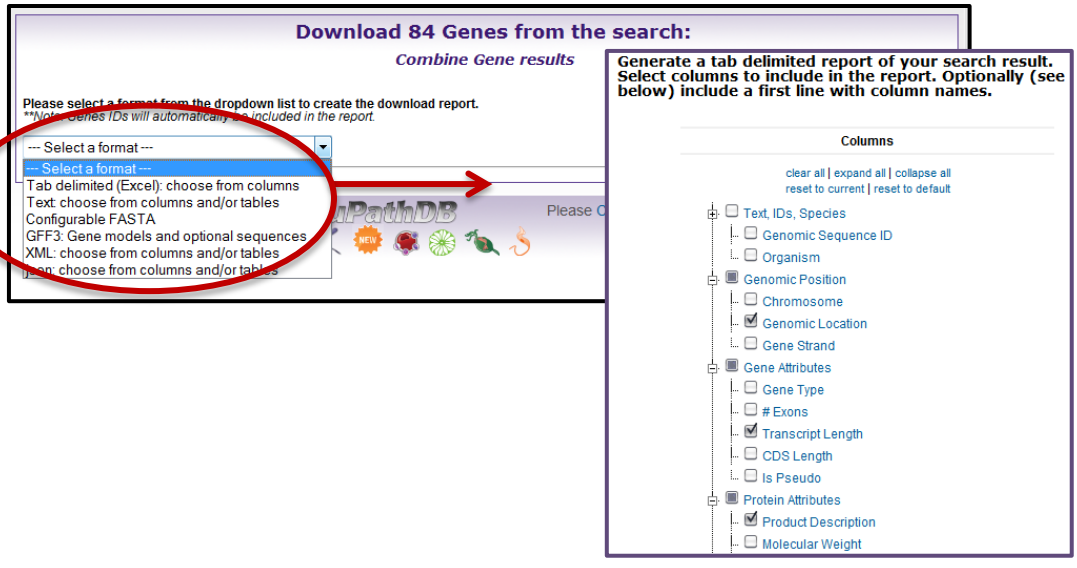
9.1 Downloading a set of results and associated data.

For this exercise you can start with any gene list of results. Start with any result list you generated this morning, such as the DNA Motif search.

Download this list of results with the following associated data: Genomic Location, Product Description, Transcript Length and Predicted GO Function.
Hint: click on the Download ## Genes link.



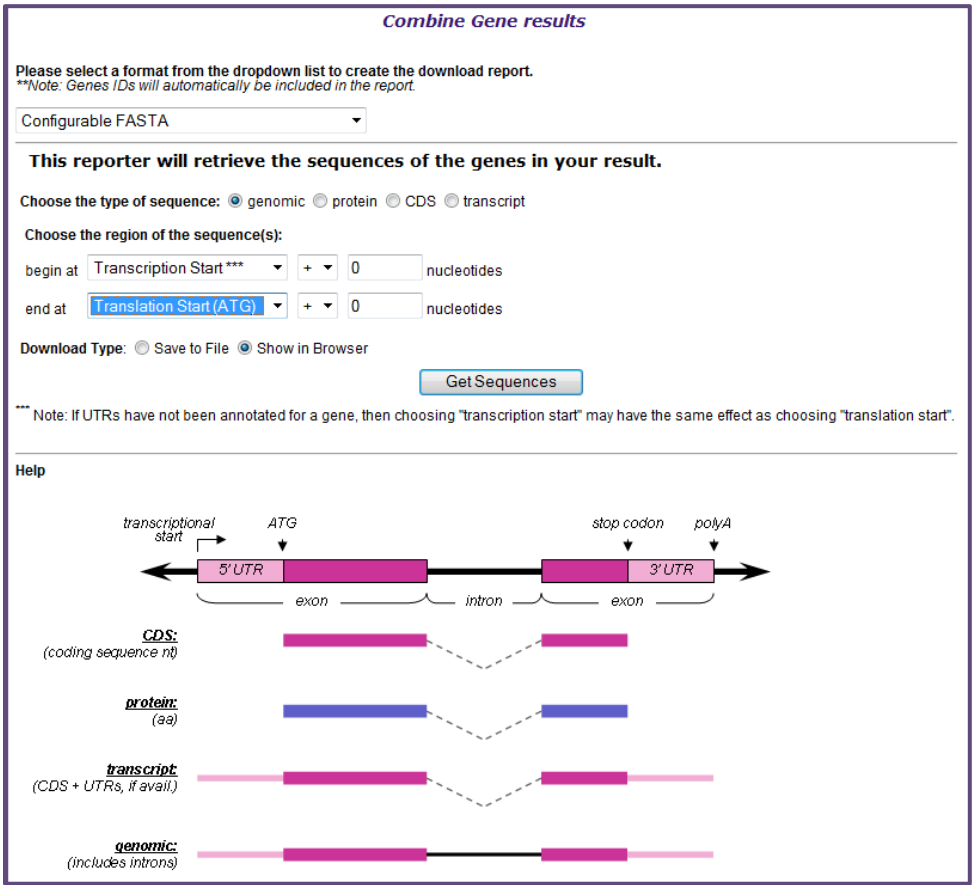
Hint: select the type of report to download and then click on the boxes to customize your report. The gene ID is automatically downloaded and so is not an option in the popup.



9.2 Download the sequences of genes in a list of results.

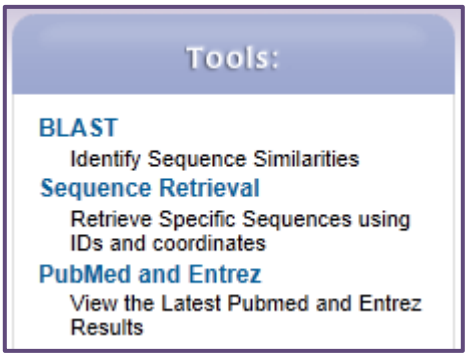
What if you are interested in examining the 5' flanking sequences of these genes? How can you easily get this sequence for subsequent analysis?

Hint: use same list of results as in 9.1. Go to the download section and select "Configurable FASTA". Now, retrieve the 500 nucleotides upstream of the start site of your genes.



Note, that you can access and download sequence with the sequence retrieval tool (SRT) accessed from the tools menu on the home page:

- Retrieve Sequences By Gene IDs.
- Retrieve Sequences By Genomic Sequence IDs.
- Retrieve Multiple Sequence Alignments by Contig / Genomic Sequence IDs.
- Retrieve Sequences By Open Reading Frame IDs.



9.3 Downloading large data files such as all coding sequences or all protein sequences for an entire genome.

Download files are available in the file download section of all EuPathDB sites
Hint: select “Data Files” under the “Download” menu in the grey tool bar.

Hint: navigate through the subfolders and find the files containing codon usage information for *T. annulata* Ankara. Folders without a strain designation contain species level data.