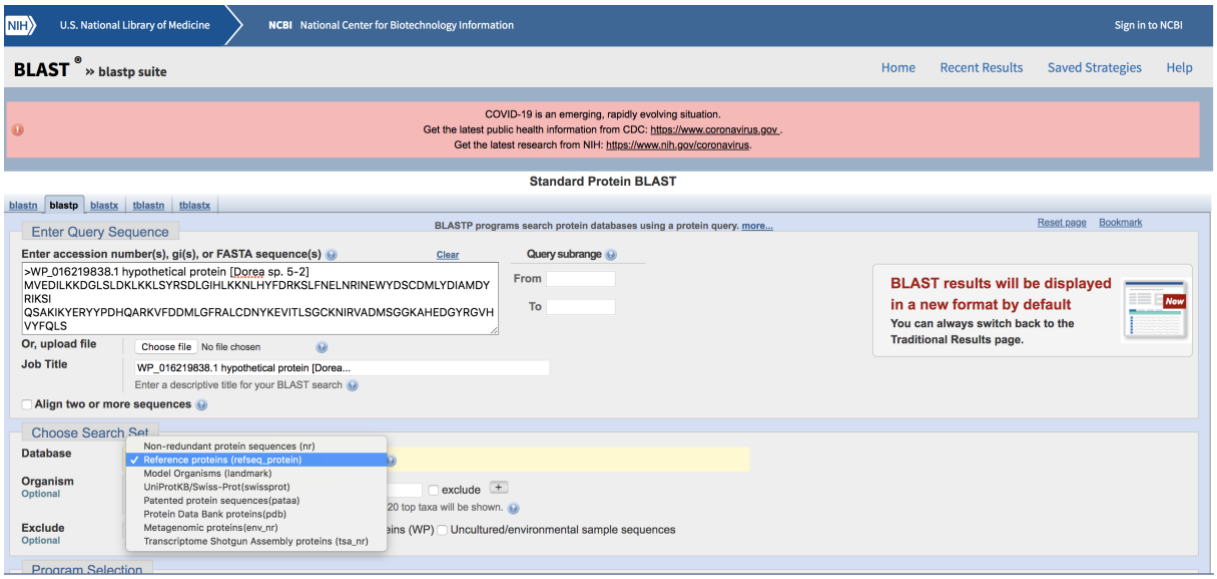
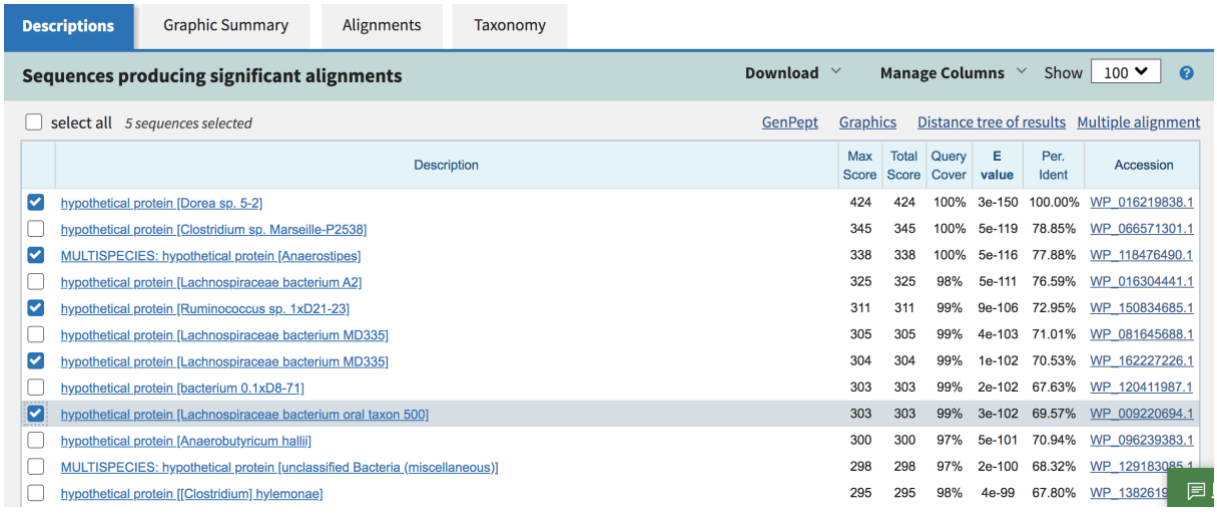
Creating an input file for FlaGs from an NCBI BlastP or PSI-Blast search

1. Run the search against RefSeq proteins:
Go to https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins , paste in your query sequence and select the refseq_protein database. Set any advanced parameters and/or organism limits you like, then click Blast.



2. When the results appear, select the proteins of interest
Use the check boxes next to each hit to select the results you're interested in, or leave "select all" checked.

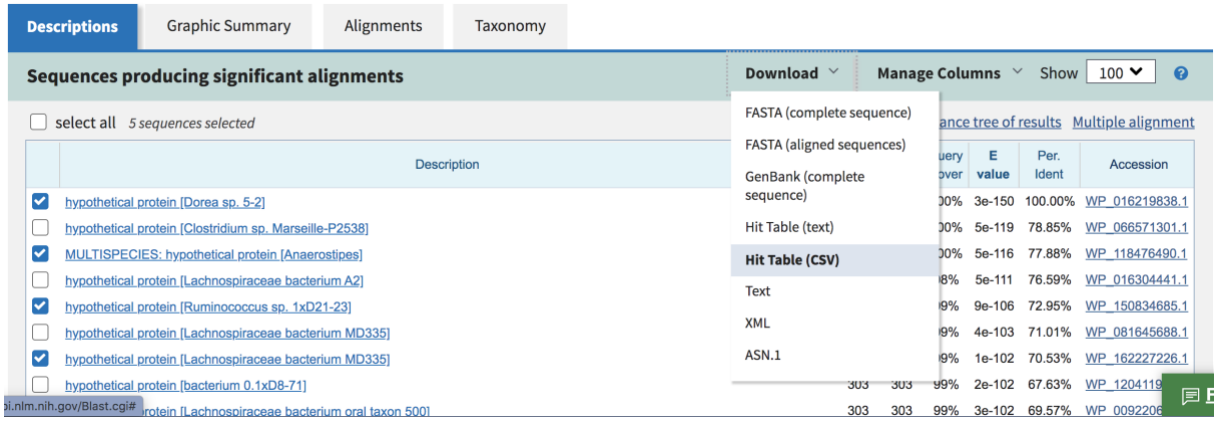


3a: If you have <100 sequences selected, you can click "Genpept"
This takes you to a multi-protein summary page. You can change the format to "Accession list" (making sure the number of results shown per page is set high enough).

The list of accessions can be pasted into a text file to use as FlaGs input. You're done and ready to run FlaGs! You can stop reading here.

```
WP_016219838.1
WP_118476490.1
WP_150834685.1
WP_162227226.1
WP_009220694.1
```

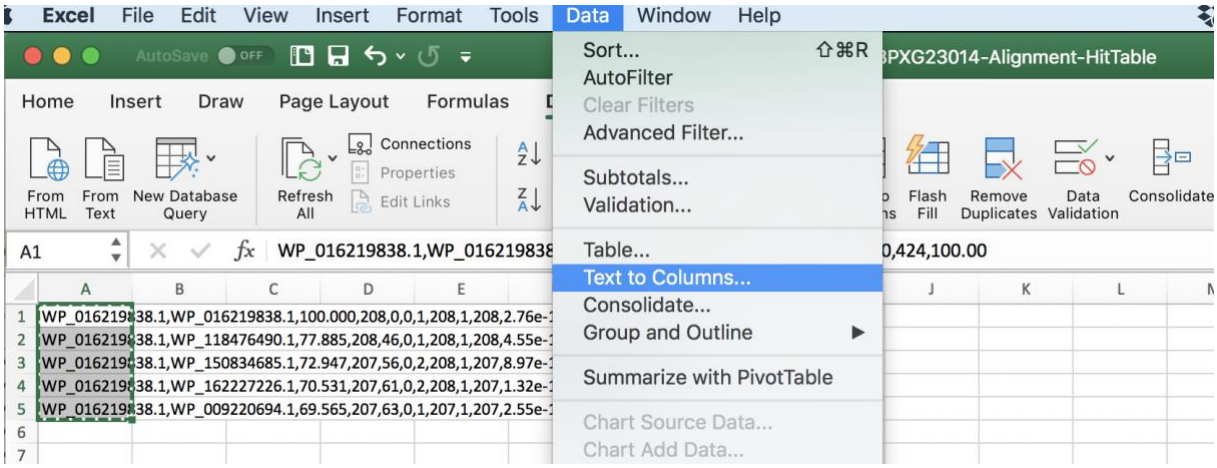
3b: If you have >100 sequences selected (or just want to try a different way), you can click "Download", then "Hit Table (CSV)"



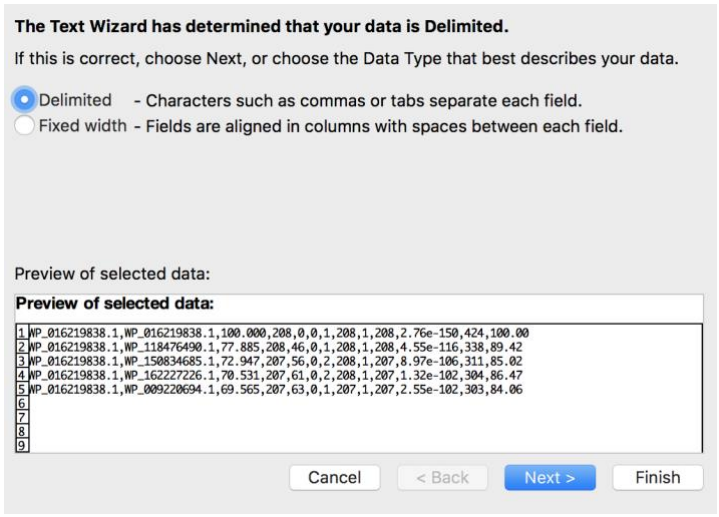
This is a table that can be opened in a spreadsheet program such as Excel, Numbers or Google Sheets

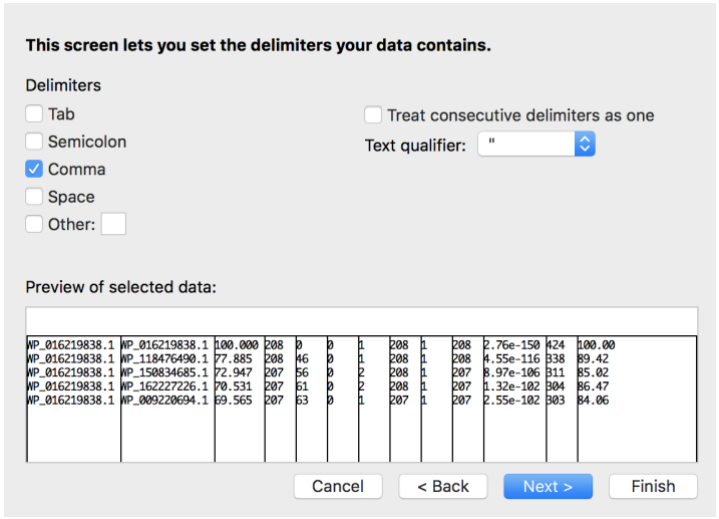
4. Open your table, and split the text into columns

In Excel and Google Sheets this is done through the Data > Text to columns menu option



Follow the wizard to split by commas





Click finish and you will see your list of accessions in the second column

The list of accessions can be copied and pasted into a text file to use as FlaGs input. You're done and ready to run FlaGs!

```
WP_016219838.1
WP_118476490.1
WP_150834685.1
WP_162227226.1
WP_009220694.1
```