Parse and Converting workshop

BLAST report

- [Ecoli proteins](#)
- [S_enterica proteins](#)
- [Ecoli-vs-Senterica.BLASTP.tab.gz](#)

Write script to read in the BLAST report. - Calculate for each alignment what the % aligned of the Ecoli query protein is? - Calculate the % of proteome that was aligned (out of the total number of Ecoli proteins)

Parsing BLAST https://github.com/biodataprog/GEN220_2019_examples/blob/master/Bioinformatics_2/parse_blast.py

Orthofinder parsing

Here is a data file - [OrthoFinder result](#)

- Let's write a script which will summarize the data by Species.

The data look like `ORTHOGROUP GENENAME_SP1, GENENAME2_SP1 GENENAME_SP2`