Exploring Proteomics Data
(Exercise 4)

1. Find all genes with mass spec evidence from the basal body in Giardia. (For this exercise use http://giardiadb.org)

- The proteomics search is available under the heading “Protein Expression” in the “Identify gene by” section.

- How many genes did you identify? Examine the filter table below the strategy section. Why is the number in the ortholog groups column less than the number in the Assemblage A column? Why do the other cells have a zero?
- What kinds of genes are in your result list? (hint: one option is to use the word cloud tool - click on the little graphic icon next to the column called “Product description”).

- How many genes with mass spec data from the basal proteome experiment also have mass spec evidence from the “mitosome enriched fraction experiment”?
- The default parameters of the mass spec search is to identify any gene with at least 1 peptide identified. How will your results change if you revise the two steps in your search strategy to only return genes with at least 5 peptides identified?

2. Find gene in Cryptosporidium that have mass spec evidence from any of the sporozoite proteomics experiments available in CryptoDB. (For this exercise use http://cryptodb.org)

- Explore the available proteomics data and select samples that make sense. You may need to click on the ‘+’ sign to expand experiments to see the underlying samples.
- How can you remove any gene with peptide evidence from non-sporozoite samples? (hint: add a step for mass spec data and think about how you will combine your results.)

![Add Step 2: Mass Spec. Evidence](image)

- How many of these genes are also predicted to be secreted?

![Combine Genes in Step 1 with Genes in Step 2](image)

- Note that so far you have been searching for C. parvum genes because we only have proteomics data from this species. However, what if you are interested in the orthologs of these genes in C. muris. How can you transform your C. parvum results to C. muris genes? (hint: add a step then select the “transform by orthology” option.)

![Add Step 4: Transform by Orthology](image)

- Why did the number of C. parvum genes increase?