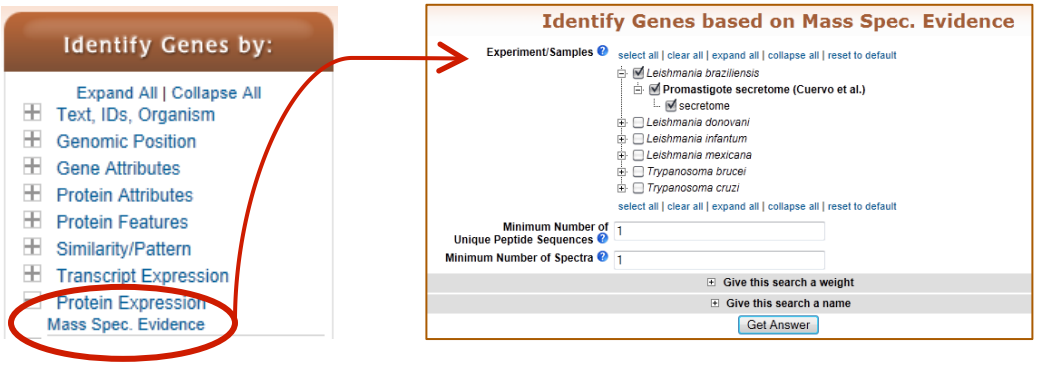
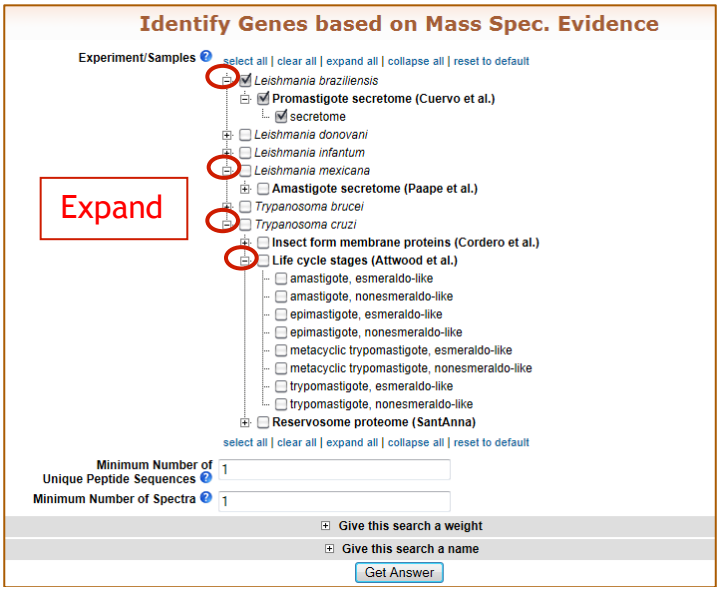
Exploring Proteomics Evidence
Exercise 12

11.1 Exploring proteomics data in TriTrypDB.

a. How many organisms have mass spectrometry evidence in TriTrypDB?
Hint: take a look at the "Mass Spec. Evidence section", under protein expression.



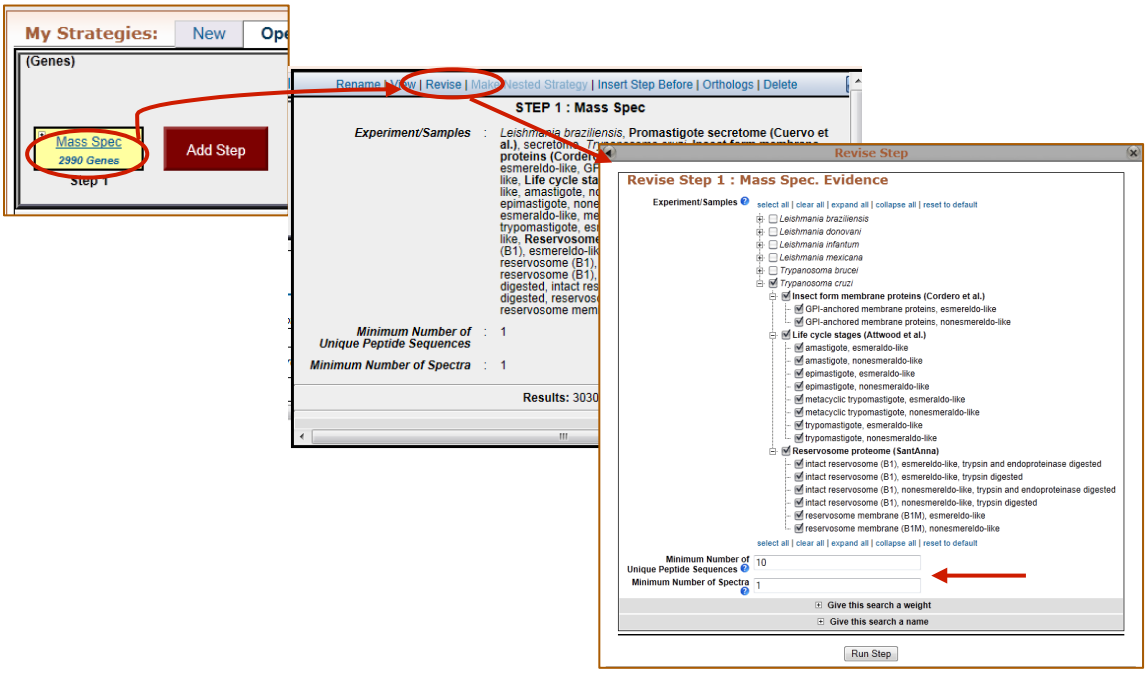
b. What kinds of experiments and parasite stages are represented?
Hint: expand species and experiments by clicking on the plus signs.



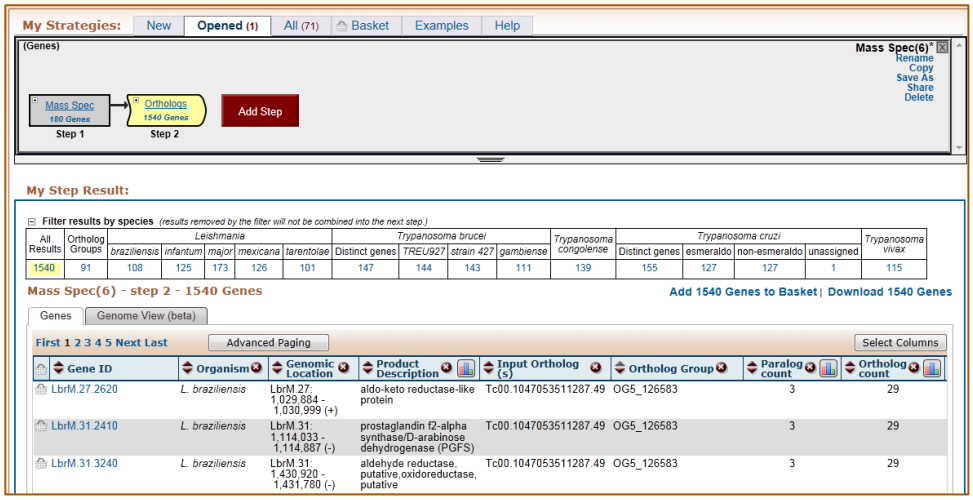
11.2 Finding all genes with mass spec evidence in *T. cruzi*.

a. How many genes in *T. cruzi* have mass spec evidence?
Hint: select *Trypanosoma cruzi* from the Mass Spec experiment list you explored in 10.1.

b. **How many genes from the results in a. have at least 10 unique peptide hits?** Hint: try revising the step in 'a' and change the "minimum number of unique peptide sequences" option to 10.



c. **Can you expand the list of results in 'b' to include possible paralogs in *T. cruzi*?**
Hint: you will have to use the ortholog transform option when adding a step and select only *T. cruzi*. Explore the columns in your result set.

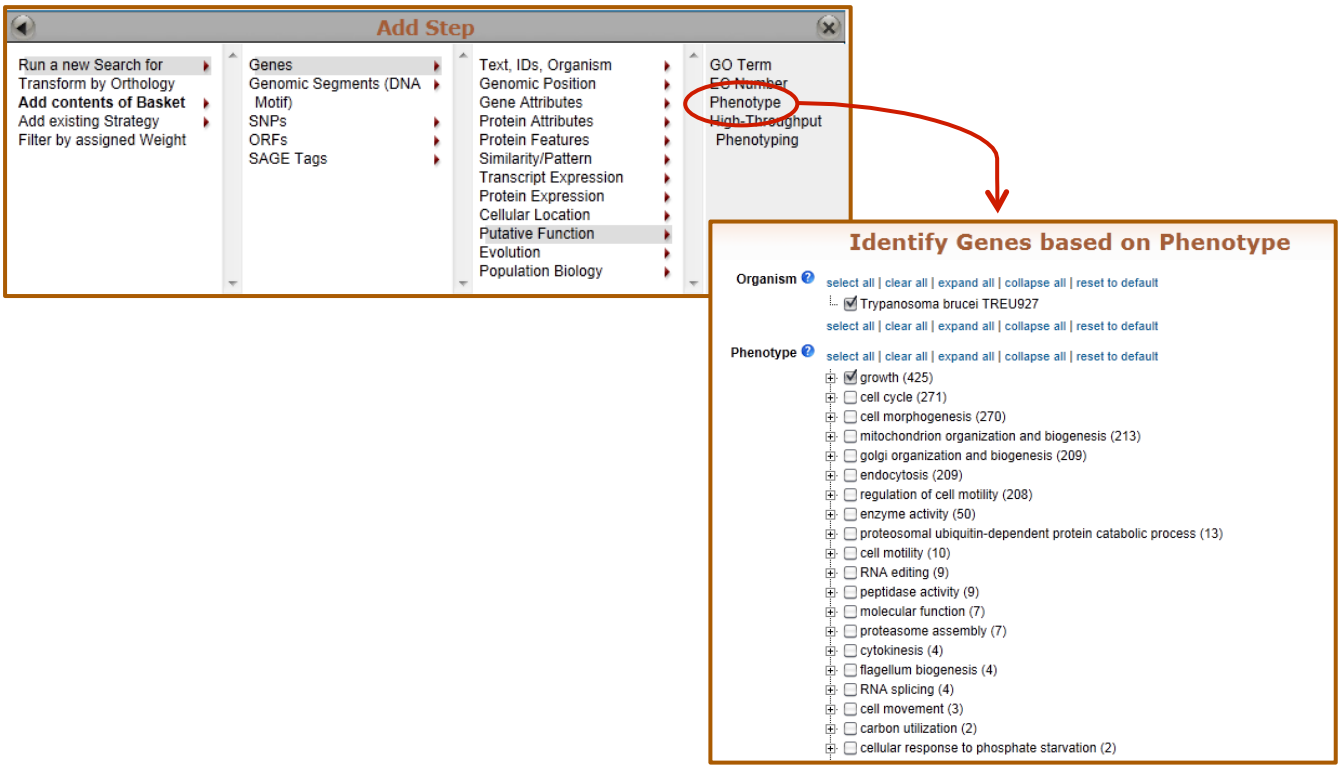


11.3 Proteins with post-translational modifications.

a. **Find all genes whose proteins have evidence of post-translational modification in *L. donovani*.** How many did you get?

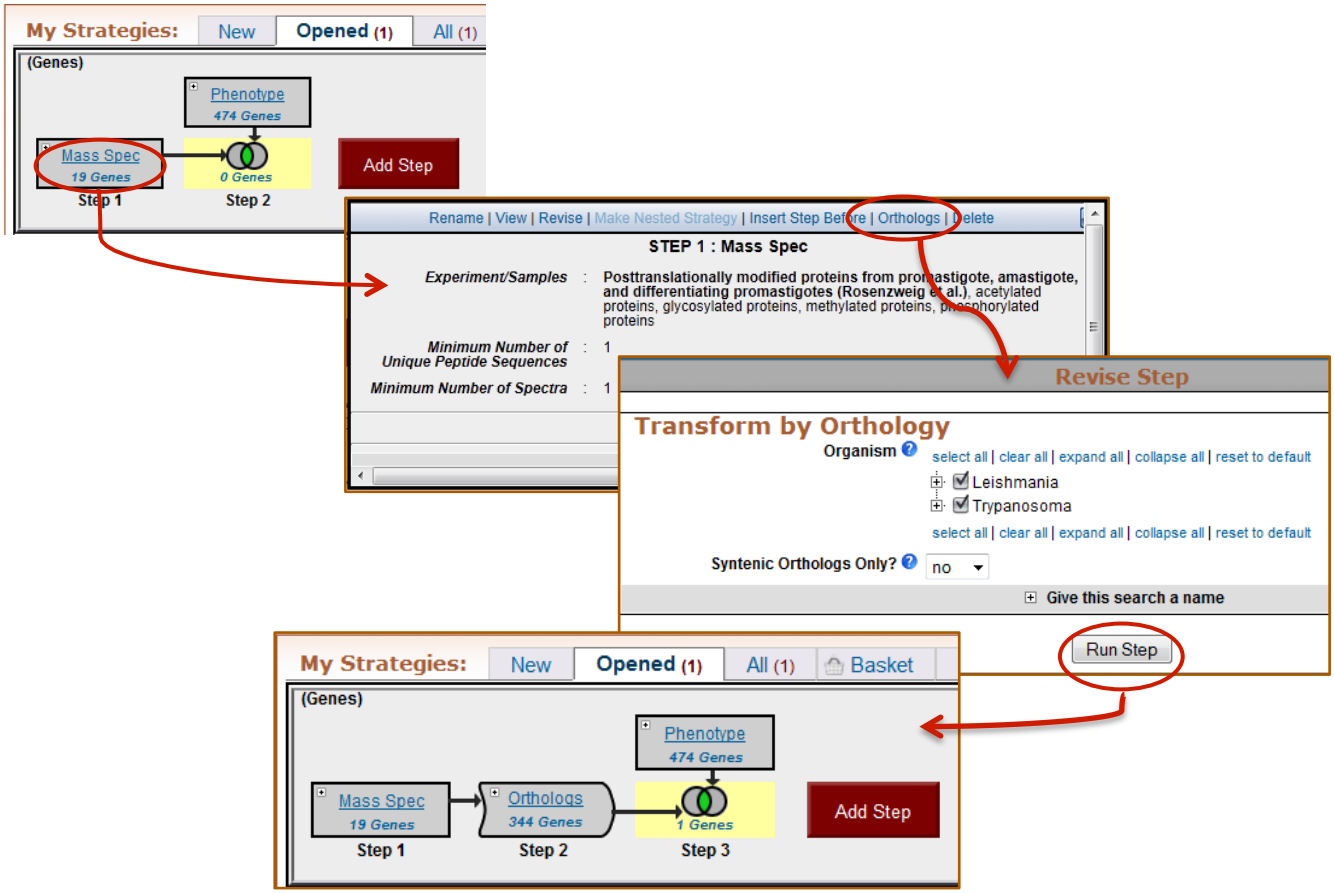
b. **How many have evidence of phosphorylation?**
Hint: revise your step from 'a' and select only the phosphorylation option.

c. **How many of these have any phenotypic evidence?**
Hint: add a step for phenotype found under "Putative Function".

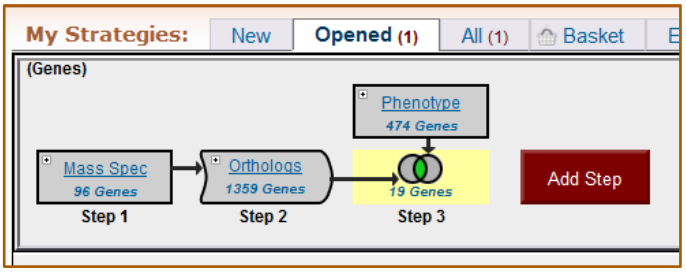


d. **How many results did you get?** Did you get zero results? Why?
Hint: where does the phenotype data come from, which organism?

Is there anything that can be done to get some results?
Hint: how about finding the orthologs of the *L. donovani* genes in other kinetoplastida.



e. **How many genes did you get?** What happens if you revise your first step to include all types of post-translational modifications?



11.4 Finding all genes with mass spec evidence in *L. infantum.*

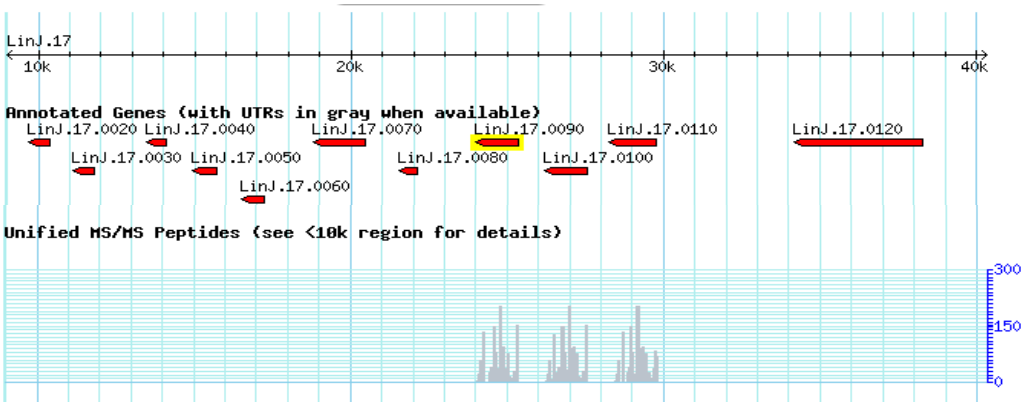
a. **Find all genes that have mass spec evidence in *L. infantum*.** How many genes did you get?

b. Which gene has the largest number of peptide hits?
Hint: sort “Number of Peptide Sequences” column.

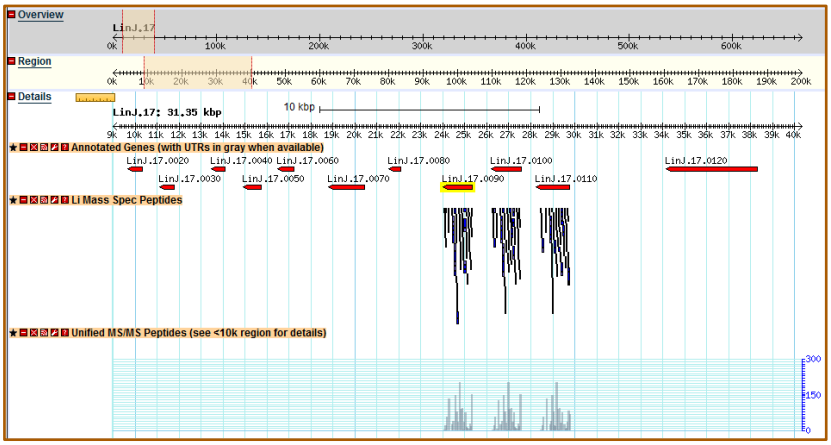
c. Which gene has the largest number of spectra?

- Did the gene with the largest number of peptides also have the largest number of spectra?
- Is this surprising or plausible?

d. **Go to the gene page for one of the genes with the largest number of peptide hits (from step b).** Take a look at the “Unified MS/MS Peptides” track in the genomic context view. What is this graphic telling you?

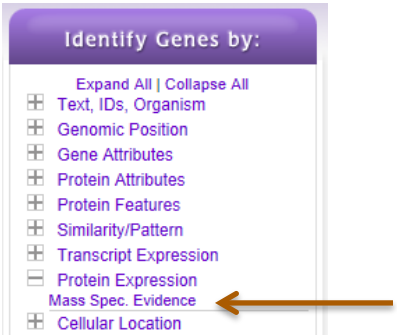


e. **View this gene in GBrowse.** Turn on the tracks for L. infantum MS/MS Peptides and for Unified MS/MS Peptides. Do you see a correlation between the graph and the peptides?
Hint: you may wish to turn off all other tracks to make it easier to visualize.



11.5 Finding genes with mass spec evidence in *P. berghei* gametocytes.
Note: For this exercise use http://www.plasmodb.org

a. **Find all *P. berghei* genes that have mass spec evidence in either or both male and female gametocytes.**
Hint: mass spec searches are in the "protein expression" expression section. Either or both is the Union of both results, not the intersection.



- How many genes did you get? How did you get to this number?
- Try running this search in two different ways:
  i. Select both male and female gametocyte options and run the search.
  ii. Select one of them first, run the search then add the other one using the add step button. How did you combine the two steps? Do you get the same results as in (i)?

b. **Find all genes that have mass spec evidence in both male and female gametocytes.**
Hint: use the strategy you developed in (ii) to get this answer, but change the union into an intersection.

c. **Find genes that have mass spec evidence only in male gametocytes and not in female ones**.
Hint: modify the set operation in b.

d. Find genes that have mass spec evidence only in female gametocytes and not in male ones.
Hint: modify the set operation in b.

e. **Which female gametocyte gene has the highest number of peptide sequences?**
Hint: look at the "number of peptide sequences" column in the list of results.

f. **What does the distribution of peptides in the gene from 'e' look like?**
Hint: go to the gene page and look at the "Protein features" section, or go to the genome browser from the gene page and turn on the right tracks.