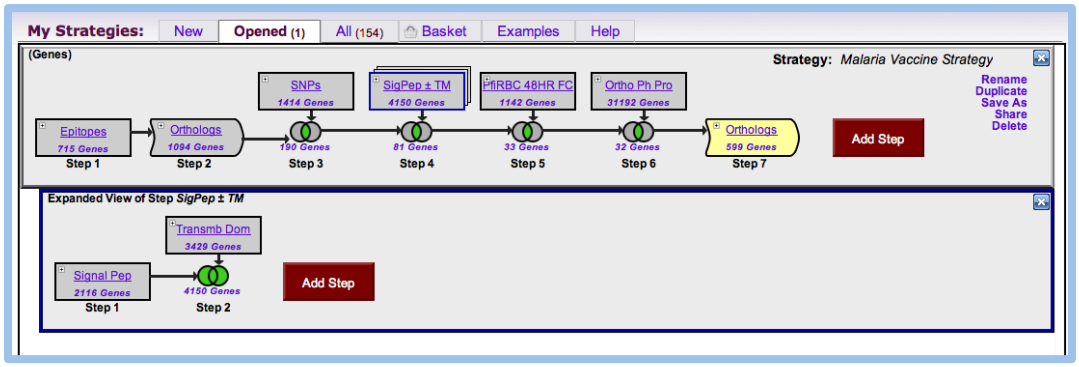
Data Integration II
Complex strategy: finding vaccines and drug targets
Exercise 11

11.1 Defining possible vaccine candidates in malaria.
Note: for this exercise use http://plasmodb.org

a. Considering all of the many 'Queries & Tools' available on PlasmoDB, how many criteria can you define that might be useful for identifying candidate vaccine targets?

b. After *first* trying to develop your own query, you might be interested to look at an example query:

Malaria Vaccine Strategy - http://plasmodb.org/plasmo/im.do?s=14627aa18a8052a5
Note: copy and paste the URL into your browser if the link does not work.

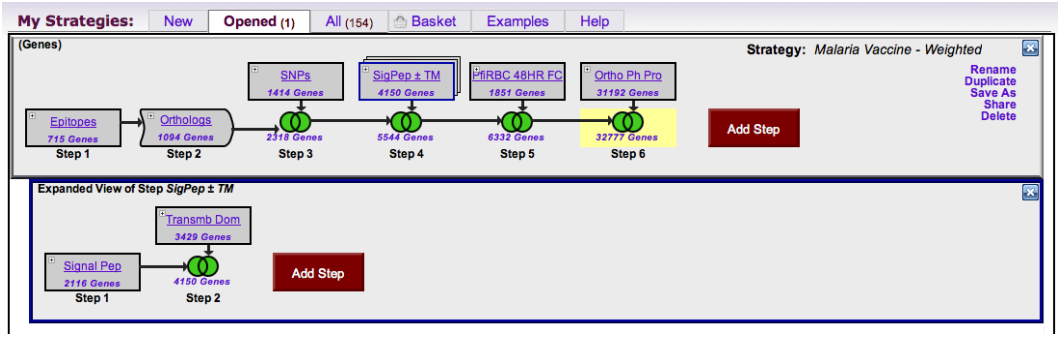


Try revising various components of this query to improve it still further to reflect your own insights, theories or experience. ***Note that if you have logged in, you can save the results of your queries for future reference, or to share with others.***

c. How would your results change if you used weighted searches? Revise each step of your strategy and assign a weight to it. The weight is arbitrary; you decide on the scale and the results are sorted based on the sum of the weights. Remember, for weighting to work you have to use the Union operation to join the steps. After doing this, what are some of your top candidates? (Hint: you can sort the results columns by clicking on the arrows next to the column name.) Here is the example from above with assigned weights (note that the weighted strategy may take a while to load):

Malaria Vaccine – Weighted - http://plasmodb.org/plasmo/im.do?s=094f2aae5385bbd9

Note: copy and paste the URL into your browser if the link does not work.



11.2 Defining possible drug targets in Trypanosomes.
Note: for this exercise use http://tritrypdb.org

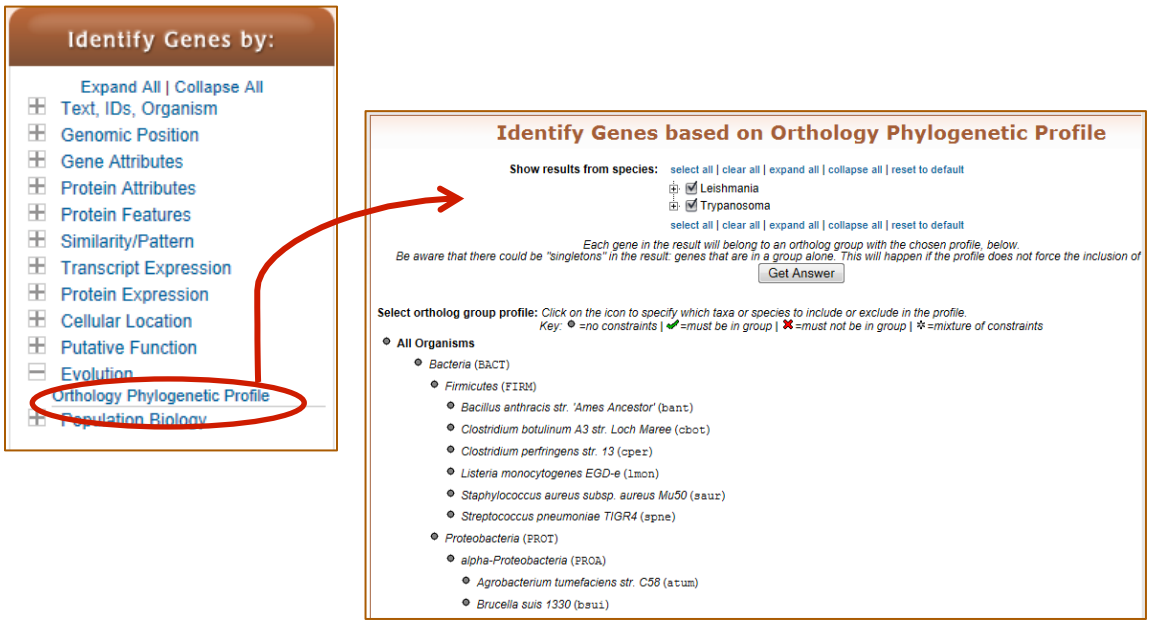
a. Consider your ideal drug target – what may be some of its characteristics?

- Would it be useful if the drug target is an enzyme?

- Would it be better to identify something conserved between the parasite and the host or not?

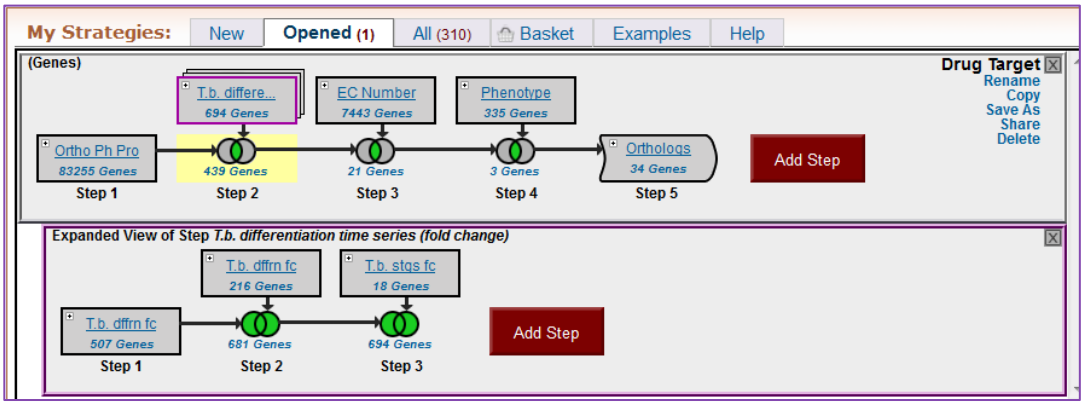
- What about the biology of the parasite you are working on? When would you prefer this protein to be expressed? For example, in *T. brucei* an enzyme may be more active during the slender replicative form. Alternatively, you may decide to concentrate on proteins involved in differentiation of the parasite. For example, this may include proteins that are differentially expressed between slender and the stumpy (cell cycle arrested) forms of *T. brucei*.

b. How would you build your strategy? One place to start is to ask for all proteins in TriTrypDB that do not have orthologs in mammals. You may wish to add a step for anything with an enzyme commission (EC) number. Using microarray evidence to specifically look for differentially regulated genes between slender and stumpy forms might be useful (hint: use a nested strategy to combine multiple microarray experiments). What about other data types, for example, phenotype data?



c. Here is a sample strategy that you may wish to consult. Copy and paste the URL into your browser if the hyperlink does not work.

http://tritrypdb.org/tritrypdb/im.do?s=e1d4776be3f9b558



d. What would happen if you added weights to your strategy? Give it a try! Here is the above strategy with weights added. Copy and paste the link into your browser if the hyperlink does not work.

http://tritrypdb.org/tritrypdb/im.do?s=ff340e3a1963dec9