

Complex strategy: finding vaccines and drug targets

The exercises below are meant as example complex strategies to illustrate how various data types may be combined.

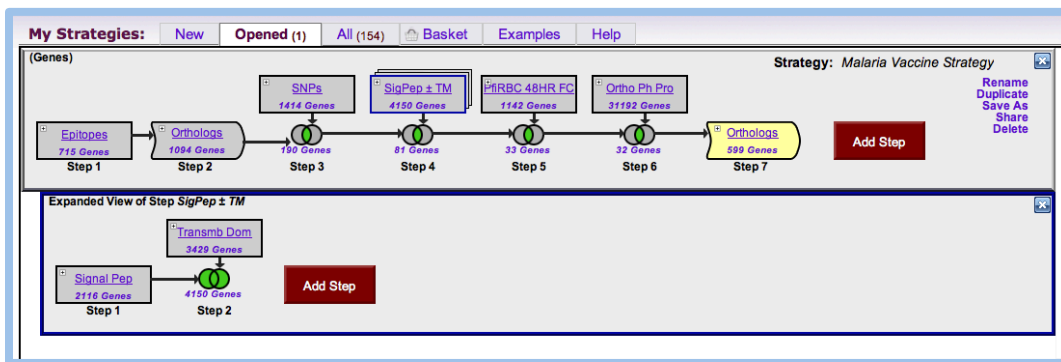
1. Defining possible vaccine candidates in malaria.

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.



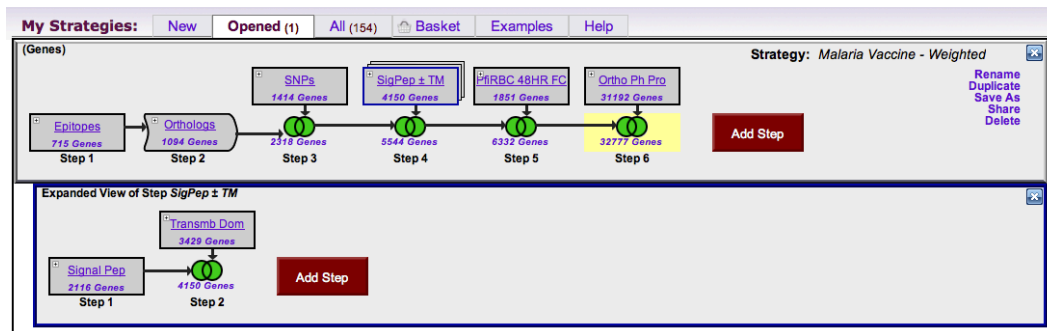
Try revising various components of this query to improve it still further to reflect your own insights, theories or experience. **Note that if you have logged in, you can save the results of your queries for future reference, or to share with others.**

- 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.



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?

Identify Genes by:

- Expand All | Collapse All
- Text, IDs, Organism
- Genomic Position
- Gene Attributes
- Protein Attributes
- Protein Features
- Similarity/Pattern
- Transcript Expression
- Protein Expression
- Cellular Location
- Putative Function
- Evolution
- Orthology Phylogenetic Profile**
- Population Biology

Identify Genes based on Orthology Phylogenetic Profile

Show results from species: Leishmania Trypanosoma

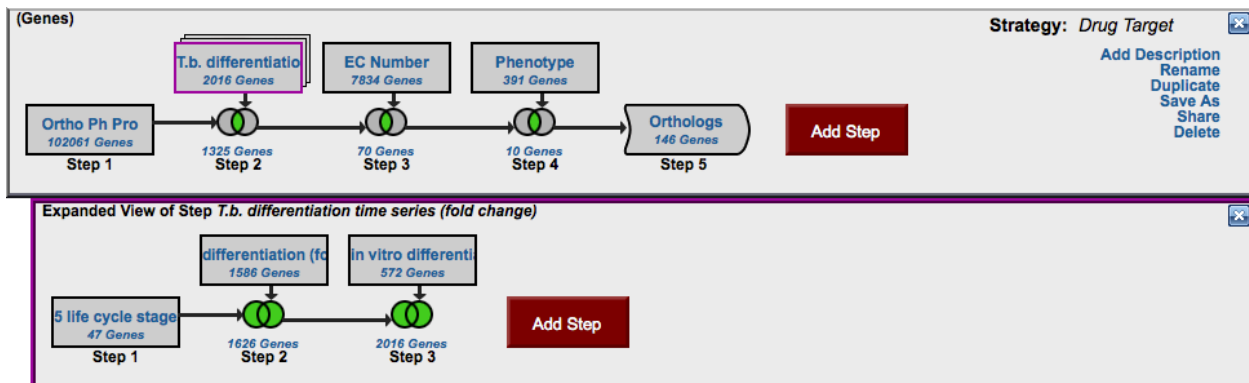
Each gene in the result will belong to an ortholog group with the chosen profile, below. Be aware that there could be "singletons" in the result: genes that are in a group alone. This will happen if the profile does not force the inclusion of

Select ortholog group profile: Click on the icon to specify which taxa or species to include or exclude in the profile.
Key: =no constraints | =must be in group | =must not be in group | =mixture of constraints

- All Organisms
 - Bacteria (BACT)
 - Firmicutes (FIRM)
 - Bacillus anthracis str. Ames Ancestor (bant)
 - Clostridium botulinum A3 str. Loch Maree (cbot)
 - Clostridium perfringens str. 13 (cpez)
 - Listeria monocytogenes EGD-e (lmon)
 - Staphylococcus aureus subsp. aureus Mu50 (saux)
 - Streptococcus pneumoniae TIGR4 (spne)
 - Proteobacteria (PROT)
 - alpha-Proteobacteria (PROA)
 - Agrobacterium tumefaciens str. C58 (atum)
 - Brucella suis 1330 (bsui)

- 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://tritypdb.org/tritypdb/im.do?s=ff340e3a1963dec9>

My Strategies: [New](#) [Opened \(2\)](#) [All \(410\)](#) [Basket](#) [Public Strategies \(5\)](#) [Help](#)

(Genes) **Strategy: Drug Target -- Weighted**

Ortho Ph Pro
102061 Genes
Step 1
T.b. differentiatio
1996 Genes
Step 2
EC Number
7834 Genes
Step 3
Phenotype
391 Genes
Step 4
Add Step

5 life cycle stage
47 Genes
Step 1
differentiation (fc)
1588 Genes
Step 2
in vitro differenti
542 Genes
Step 3
Add Step

107468 Genes from Step 4
Strategy: Drug Target -- Weighted [Add 107468 Genes to Basket](#) | [Download 107468 Genes](#)

Click on a number in this table to limit/filter your results

Gene Results [Genome View](#)

First 1 2 3 4 5 Next Last [Add Columns](#)

Gene ID	Genomic Location	Product Description	Search Weight
Tb927.9.5900	Tb927_09_v5.1: 995,071 - 998,285 (+)	glutamate dehydrogenase (GDH)	290
Tb927.6.510	Tb927_06_v5.1: 225,806 - 226,480 (+)	GPEET2 procyclin precursor,PARP A-alpha,procyclin A-alpha,procyclic form specific polypeptide A-alpha...	260
Tb927.10.5760	Tb927_10_v5.1: 1,427,705 - 1,430,174 (-)	adenylate kinase, putative	250
Tb927.1.1270	Tb927_01_v5.1: 334,231 - 335,988 (+)	homocysteine S-methyltransferase, putative	240