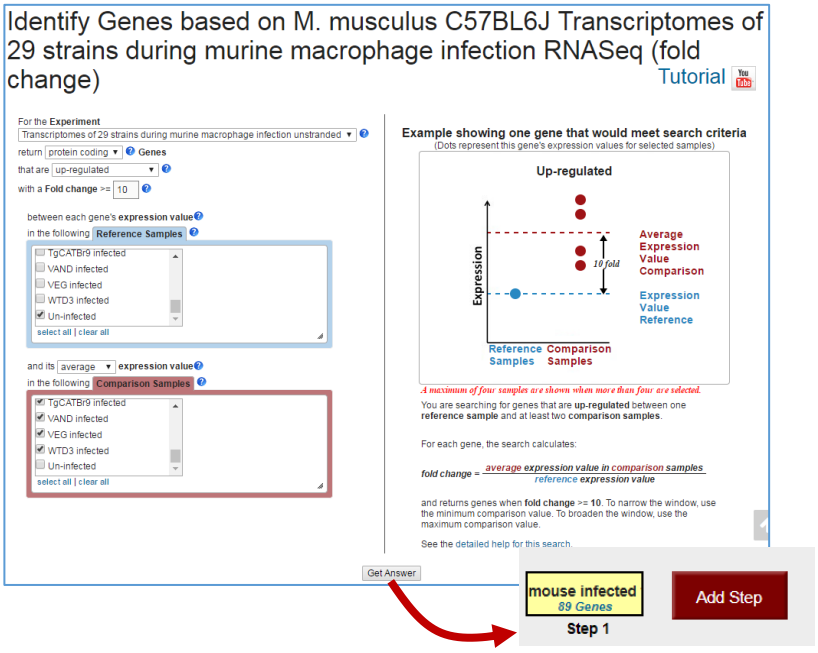
Host Response

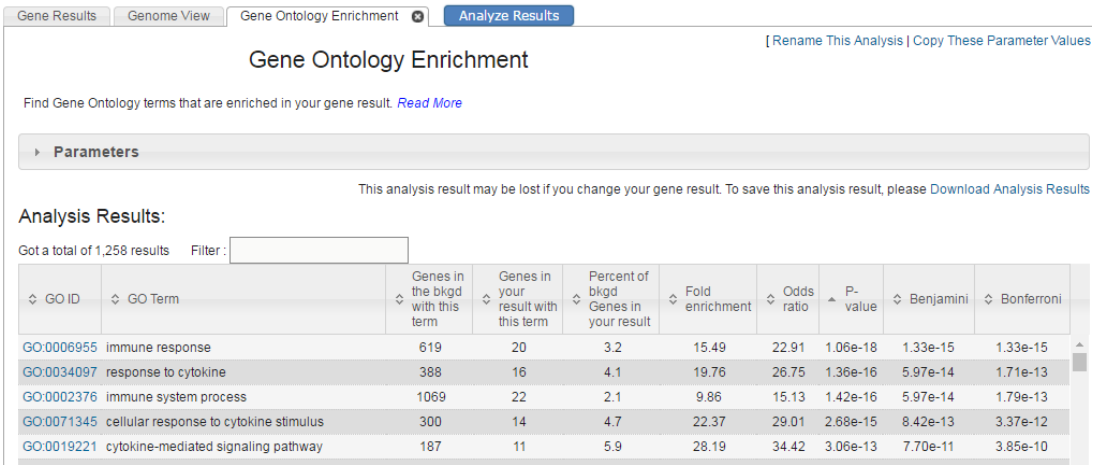
1. Find host genes that are upregulated in infected mouse cells compared to uninfected ones. For this exercise use http://hostdb.org

   a. Navigate to the "Transcriptomics" section then select "RNA Seq Evidence". Select the fold change query for the "Transcriptomes of 29 strains during murine macrophage infection (Minot et al.)" experiment.

   b. Configure the search to return genes that are up-regulated at least 10-fold compared to the uninfected control. Make sure to select upregulated. In the example below a fold change of 10 was selected and the "average" operation was applied on the comparison samples.



c. What are the functional characteristics of the genes in this result? What kinds of GO terms are enriched? Does the host immune response appear to be turned on?
Hint: click on the "Analyze Results" tab and perform a GO enrichment analysis for the biological process ontology.

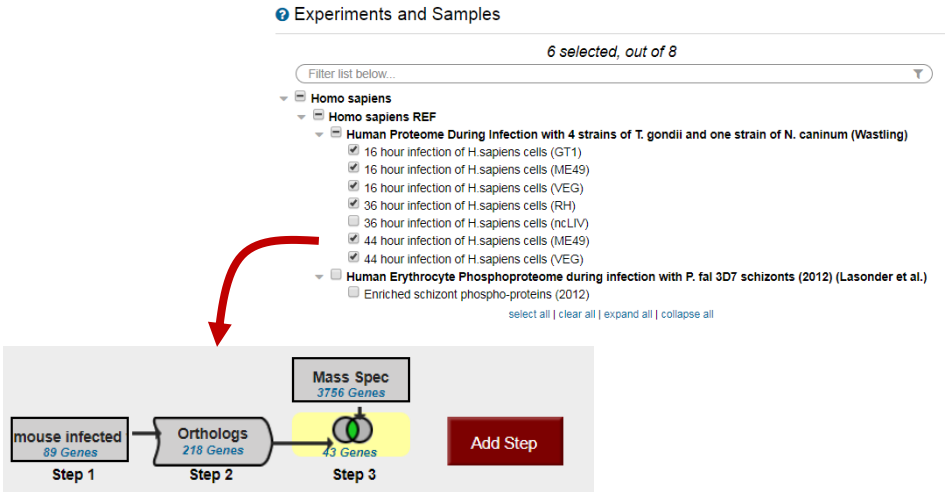


GO ID	GO Term	Genes in the bkgd with this term	Genes in your result with this term	Percent of bkgd Genes in your result	Fold enrichment	Odds ratio	P-value	Benjamini	Bonferroni
GO:0006955	immune response	619	20	3.2	15.49	22.91	1.06e-18	1.33e-15	1.33e-15
GO:0034097	response to cytokine	388	16	4.1	19.76	26.75	1.36e-16	5.97e-14	1.71e-13
GO:0002376	immune system process	1069	22	2.1	9.86	15.13	1.42e-16	5.97e-14	1.79e-13
GO:0071345	cellular response to cytokine stimulus	300	14	4.7	22.37	29.01	2.68e-15	8.42e-13	3.37e-12
GO:0019221	cytokine-mediated signaling pathway	187	11	5.9	28.19	34.42	3.06e-13	7.70e-11	3.85e-10

d. Expand the result set to include human orthologs/paralogs of these genes. *Hint:* add a "Transform by Orthology" step choosing Homo sapiens.

e. Does this set of human genes also show enriched GO terms? What, if any, are the enriched GO terms?

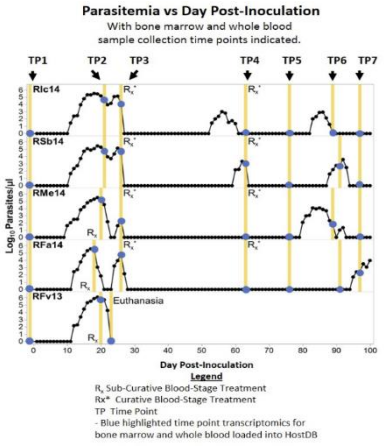
f. Do any of these human genes also have peptide evidence for their expression during infection? *Hint:* add a step and explore the proteomics data "Human Proteome During T. gondii infection"



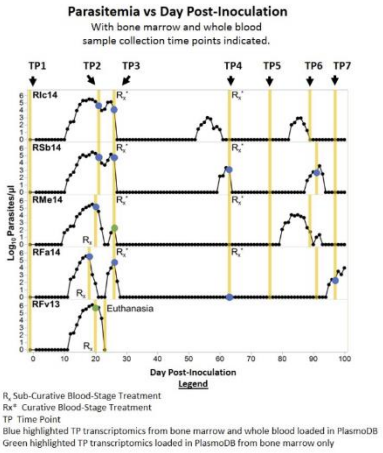
2. **Host vs parasite expression changes in dual RNA seq samples with clinical data. -** The Malaria Host-Pathogen Interacting Center (MaHPIC) aims to characterize host-pathogen interactions during malaria infections. Dual RNA seq data (host and pathogen) from MaHPIC Experiment 4, a 100 day infection of *M. mulatta* with *P. cynomolgi,* is integrated into HostDB and PlasmoDB, respectively. Five monkeys were inoculated with *P. cynomolgi* sporozoites on day 1 and followed for 100 days with daily blood collections, drug treatments and 7 pathology-driven time point collections of bone marrow and blood. Supporting information including parasitemia charts and disease progression definitions is available in the data set records (links below). Clinical data is available in PlasmoDB Downloads.

HostDB - *M mulatta* infected with *P cynomolgi* over 100 days
http://hostdb.org/hostdb/app/record/dataset/DS_4585d065bf
PlasmoDB - Transcriptome of *P cynomolgi* during 100-day infection in *M mulatta*
http://plasmodb.org/plasmo/app/record/dataset/DS_40a06f276b
MaHPIC Record page http://plasmodb.org/plasmo/mahpic.jsp
MaHPIC data files in PlasmoDB http://plasmodb.org/common/downloads/MaHPIC/

HostDB



PlasmoDB



Primate ID	Baseline (Pre-infection)	Acute Primary	Post-peak	Inter-relapse	Relapse
RIc14	TP1	TP2	TP3	TP5, TP7	TP4, TP6
RSb14	TP1	TP2	TP3	TP5, TP7	TP4, TP6
RMe14	TP1	TP2	TP3	TP4, TP5, TP7	TP6
RFa14	TP1	TP2	TP3	TP4, TP5, TP6	TP7
RFv13	TP1	TP2	TP3	N/A	N/A

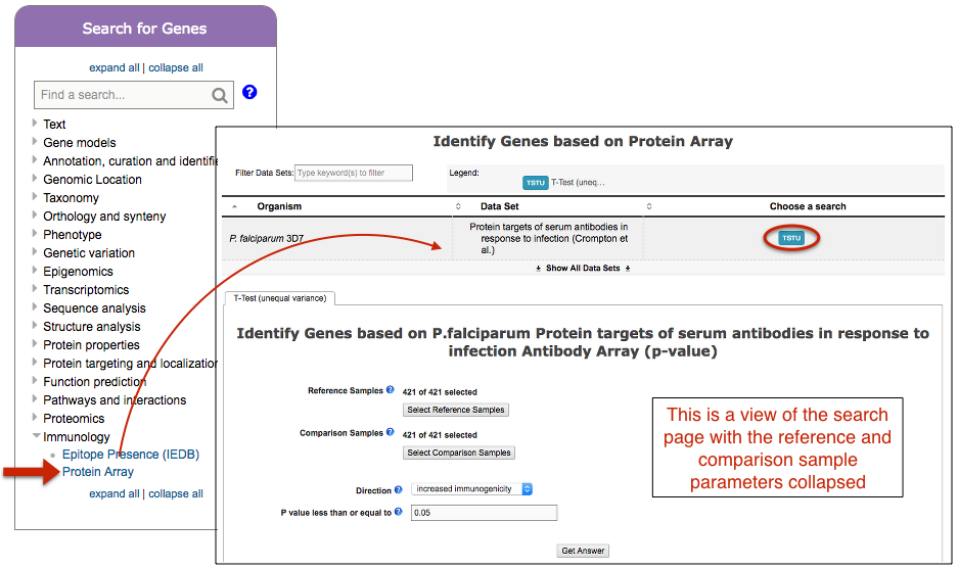
a. **Find genes that are upregulated at least 2-fold in the bone marrow of monkeys during the acute primary infection vs the post-peak infection.**

- Navigate to http://hostdb.org/hostdb/ and open the RNA sequence fold change search for the experiment called "*M mulatta* infected with *P cynomolgi* over 100 days (Joyner et al.)".
- Hint – use the information in the data records (links and small screen shots above) to create a fold change search
- Give it a try on your own ☺. Then compare your results with the strategy below or have a peak at the screenshots
  http://hostdb.org/hostdb/im.do?s=ee8250c6640668c1
  Screenshots: - https://docs.google.com/presentation/d/1oS9DUyf4hu8y9Hq95oJvhvRJ4Gh3iVekBiVCoseOghM/edit?usp=sharing
- What functional Characteristics are shared by these *M. mulatta* genes?
  Hint – run a GO enrichment for Biological process)
- What human genes are likely associated with acute primary vs post peak infections?
- What functional characteristics are shared by the human genes?
  Hint – run a GO enrichment for biological process.
- Are there differences between the GO enrichment results between the human and *M. mulatta* gene lists? If so, what underlying data contributes to the differences? What can you say about the difference between *M mulatta* genome annotation and the human genome annotation?

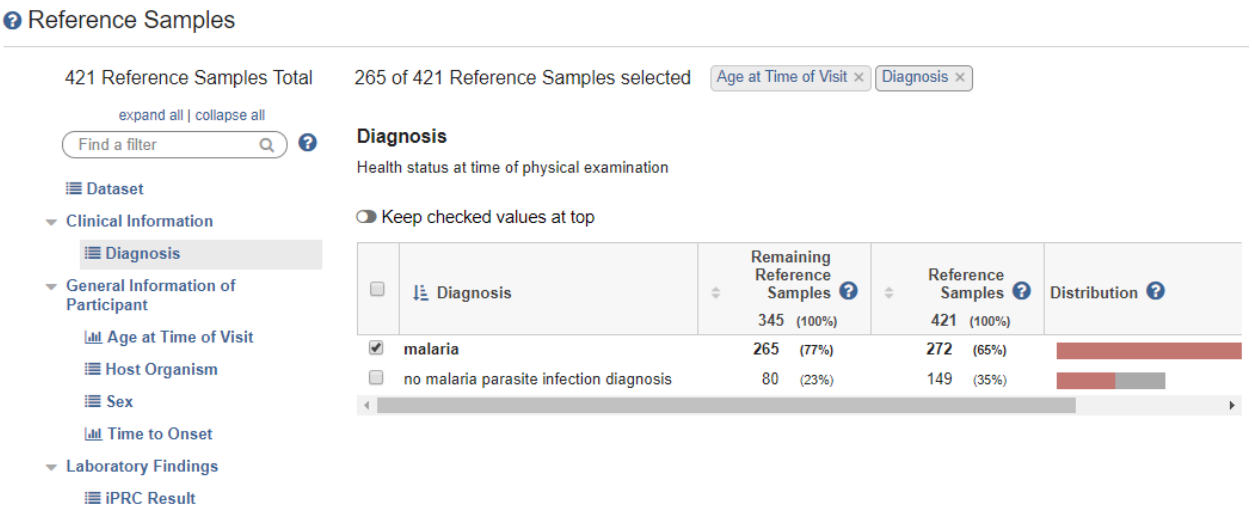
b. Create the same strategy in PlasmoDB and explore the results (minus the human transformation, of course). Do you find similar functional enrichment profiles for the plasmodium genes that are differentially expressed between acute and post-peak disease state? (Hint - Transform to an organism with mature annotation such as *P. falciparu*m before functional enrichment).

3. Find *Plasmodium falciparum* antigens that are immunogenic.
   For this exercise use http://plasmodb.org

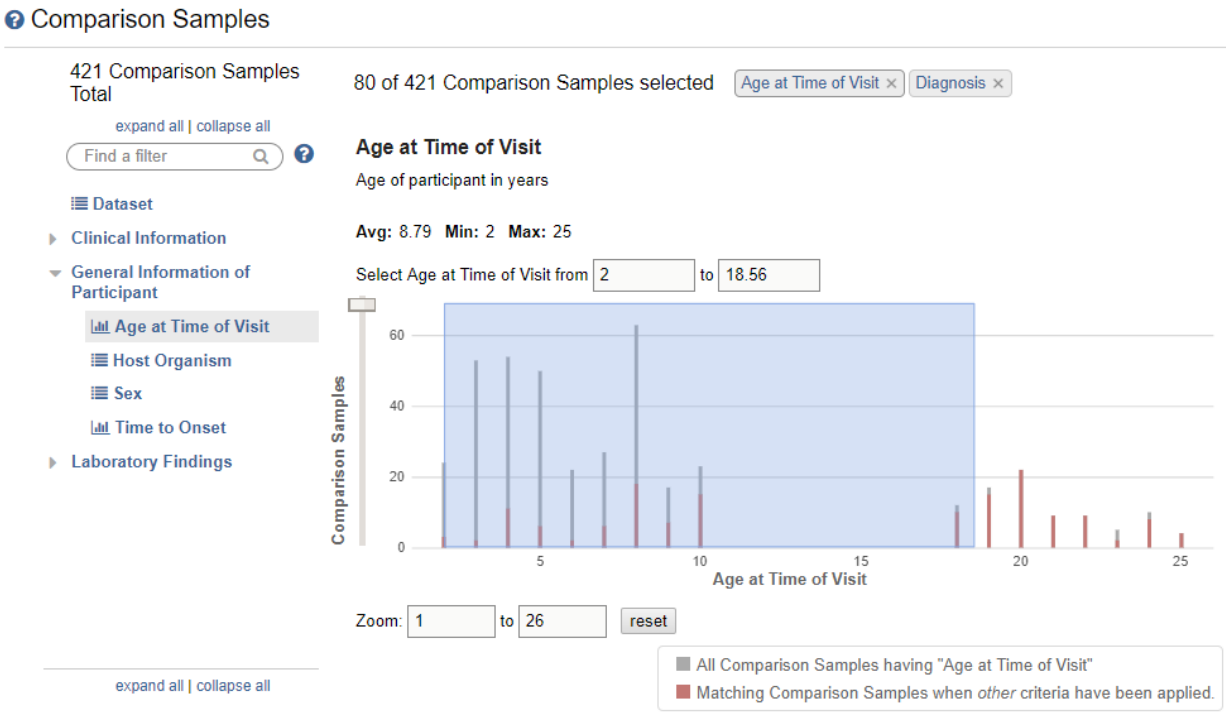
a. Identify antigens (genes) that exhibited an increased immunogenicity in children (ages 0-18) with no disease (normal) compared to children with disease (malaria). *Hint*: navigate through the “Identify Genes By”, “Immunology”, “Protein Array”. Choose the experiment "Protein targets of serum antibodies in response to infection (Crompton et al.)".



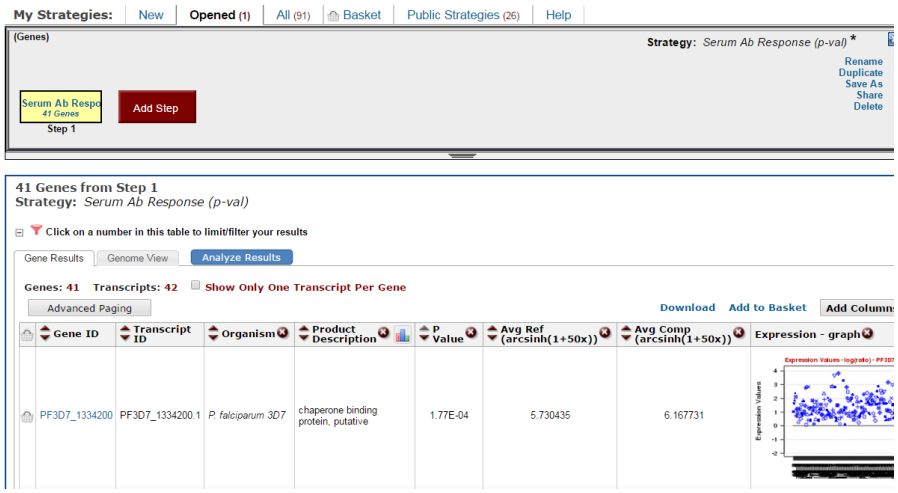
- In this example, your **reference samples will be children with a malaria diagnosis,** and your **comparison samples will be children** with **'no malaria present'**. Each set of samples (reference and comparison) has two parameters that need to be set, age and diagnosis. For the reference, set the age parameter **to 0-18** years and the Diagnosis (under Clinical Information) to **malaria**. You should be left with 265 samples in the Reference group of children with malaria.



- Move on the comparison samples and set the age to **0-18** and the diagnosis to **no malaria parasite infection diagnosis**. Your comparison group should contain 80 samples.



- The default settings for other parameters are good – increased immunogenicity and p-value = 0.05.
- You are ready to click Get Answer! What do your results look like? Could these represent potential protective antigens? (result image below)



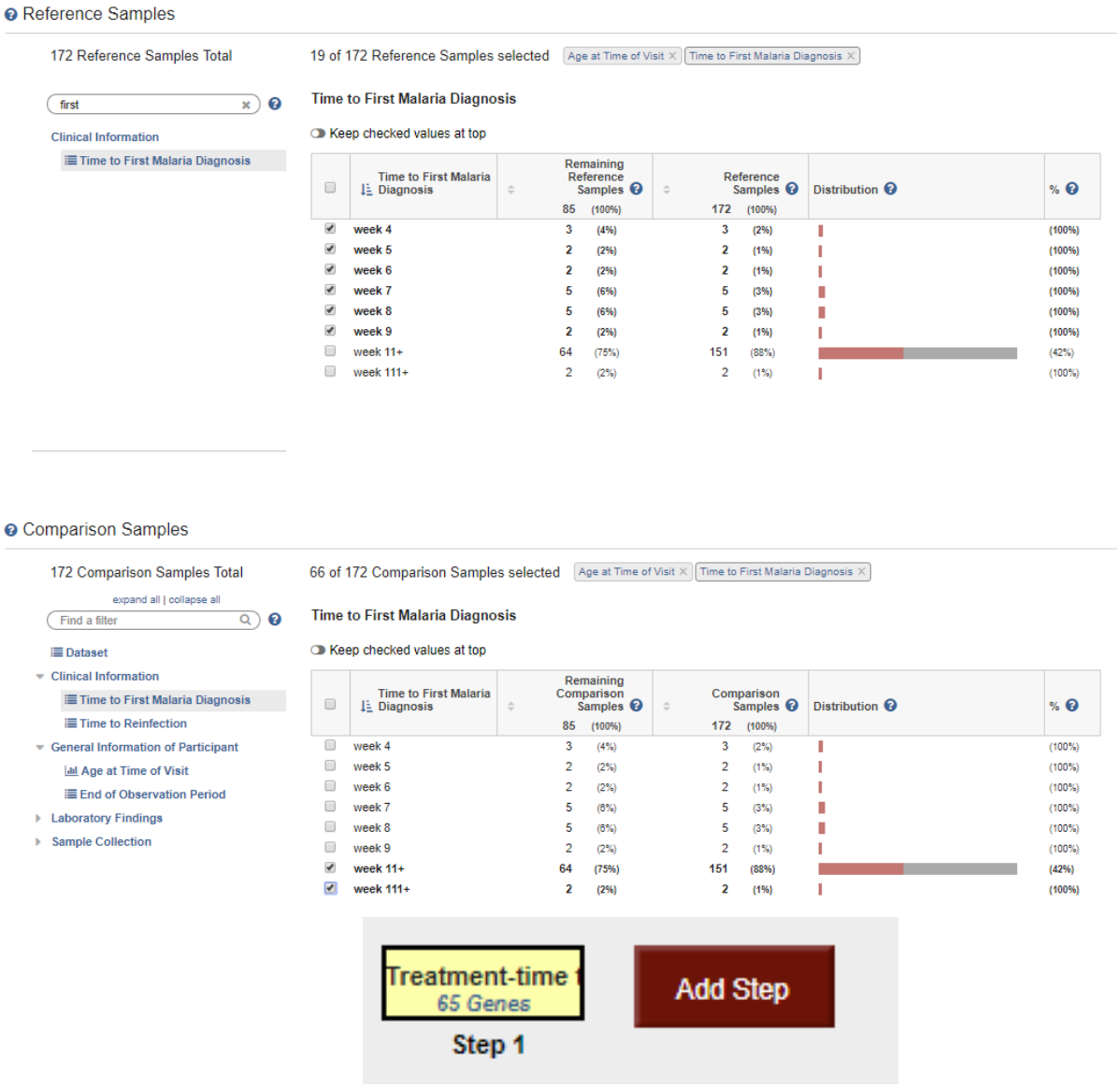
3. Find *falciparum* antigens that may be protective from reoccurrence of malaria (and potentially reinfection)

   For this exercise use http://plasmodb.org

   A recently published study from Kenya (view paper) where participants were followed for 12 weeks following an initial screening for malaria and treatment with anti-malarials is available in PlasmoDB. Each week patients were assessed for presence of parasites and clinical symptoms of malaria. Select the **"Treatment-time to reinfection cohort from Kisumu area, Kenya collected in 2003 (Dent et al.)"** experiment from the protein array searches and configure the parameters to see if you can reproduce the results of the paper. They concluded that increased antigenicity was present in children who did not show clinical symptoms of malaria, and suggest that these antigens are protective in children who did not get a recurrence of symptomatic malaria (compared to those children who did exhibit malaria symptoms). They also concluded that there did not appear to be a correlation between antigenicity and time to re-infection (could be asymptomatic). Test both these conclusions.

   Hint: compare children (age at time of visit 0-12.5) who got clinical malaria during the study (time to first malaria Dx weeks 4-9) compared to those who didn't (week 11+). Try running with increased immunogenicity then revise and change to decreased immunogenicity. See image below for help configuring the search.

   Do these results make sense?



Ask the same question (age 0-12.5) except compare time to re-infection weeks 3 and 4 with time to reinfection weeks 9,10,11,11+. Do you get significant results? Does this agree with the conclusions of the paper? Revise the search and remove the age limits, just keeping the times to re-infection.