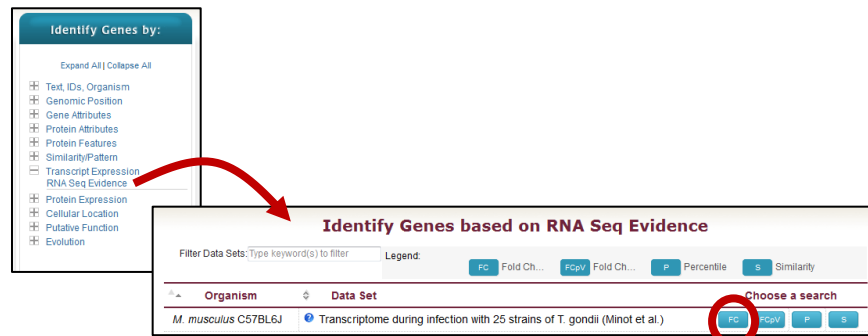


Host Response

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

- Navigate to the “Transcript Expression” section then select “RNA Seq Evidence”. Select the fold change query for the “Transcriptome during infection with 25 strains of *T. gondii* (Minot et al.)” experiment.



- Configure the search to compare all infected samples 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.

Identify Genes based on M.mus. Transcriptome during infection with 25 strains of *T. gondii* RNASeq (fold change)

[Tutorial](#) [YouTube](#)

For the Experiment
 Transcriptome during infection with 25 strains of *T. gondii* [?](#)
 return protein coding [?](#) Genes
 that are up-regulated [?](#)
 with a Fold change \geq 10 [?](#)

between each gene's expression value [?](#)
 in the following Reference Samples [?](#)

☐ T.gondii infected
☐ VAND infected
☐ VEG infected
☐ WTD3 infected
☒ not infected
 select all | clear all

and its average [?](#) expression value [?](#)
 in the following Comparison Samples [?](#)

☐ T.gondii infected
☒ VAND infected
☒ VEG infected
☒ WTD3 infected
☐ not infected
 select all | clear all

[Get Answer](#)

Example showing one gene that would meet search criteria

(Dots represent this gene's expression values for selected samples)

Up-regulated

Expression

Reference Comparison
Samples Samples

10 fold

Average Comparison

A maximum of four samples are shown when more than four are selected.
 You are searching for genes that are up-regulated between one reference sample and at least two comparison samples.

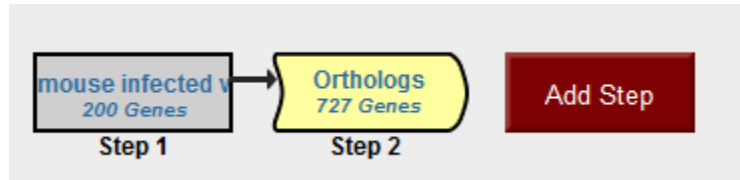
For each gene, the search calculates:

$$\text{fold change} = \frac{\text{average expression value in comparison samples}}{\text{reference expression value}}$$

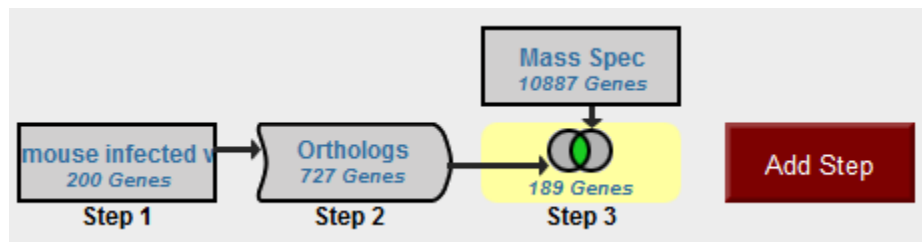
and returns genes when fold change \geq 10. To narrow the window, use the minimum comparison value. To broaden the window, use the maximum comparison value.

See the detailed help for this search.

- d. What are the human orthologs of these mouse genes? Expand the result set to include orthologs/paralogs of these genes. *Hint:* add a “Transform by Orthology” step choosing Homo sapiens.



- e. Do any of these human genes also have peptide evidence for their expression during infection? *Hint:* add a step and explore the “Mass Spec Evidence” data in the protein expression section. Run the search using the default parameters.



2. Find *Plasmodium falciparum* genes that are immunogenic in infected humans.
For this exercise use <http://plasmodb.org>

Antibodies are central to malaria immunity, which is only acquired after years of exposure to *Plasmodium falciparum* (Pf). Despite the enormous worldwide burden of malaria, the targets of protective Abs and the basis of their inefficient acquisition are unknown. PlasmoDB contains protein array data which provides a measure of a gene's 'immunogenicity' by hybridizing serum (which contains antibodies) to an array spotted with peptides from 2320 *P. falciparum* gene products. Searching this data in PlasmoDB identifies genes whose protein products acted as antigens in malaria patients.

- a. Identify genes (antigens) that exhibited an increased immunogenicity in asymptomatic children (ages 0-18) compared to children with disease (infected). *Hint:* the "Protein Array" search is available in the "Host Response" menu item in the "Identify Genes By" section of the home page. Choose the experiment **Protein targets of serum antibodies in response to infection (Crompton et al.)**.

The image shows a screenshot of the PlasmoDB website. On the left, the 'Identify Genes by:' menu is open, showing various search criteria. A red arrow points from the 'Protein Array' option in the 'Host Response' section to the 'Identify Genes based on Protein Array' search page. The search page has a table with columns 'Organism' and 'Data Set'. The first row is selected, showing 'P. falciparum 3D7' and 'Protein targets of serum antibodies in response to infection (Crompton et al.)'. A red circle highlights the 'TSTU' button next to this row. Another red arrow points from this button to the 'Identify Genes based on P.fal. Protein targets of serum antibodies in response to infection Antibody Array (p-value)' search page. This page shows search parameters: 'Reference Samples' (421 of 421 selected), 'Comparison Samples' (421 of 421 selected), 'Metadata category to color graph by' (DiseaseState), 'Direction' (increased immunogenicity), and 'P value less than or equal to' (0.05). A red text box on the right says: 'This is a view of the search page with the reference and comparison sample parameters collapsed.'

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
- Population Biology
- Host Response
 - Protein Array

Identify Genes based on Protein Array

rd(s) to filter Legend: TSTU T-Test (u...

Organism	Data Set	Choose a
P. falciparum 3D7	Protein targets of serum antibodies in response to infection (Crompton et al.)	TSTU
P. falciparum 3D7	International Centers of Excellence for Malaria Research (International Centers of Excellence for Malaria Research (ICEMR))	TSTU
P. falciparum 3D7	Treatment-time to reinfection cohort from Kisumu area, Kenya collected in 2003 (Dent et al.)	TSTU

T-Test (unequal variance)

Identify Genes based on P.fal. Protein targets of serum antibodies in response to infection Antibody Array (p-value)

Reference Samples 421 of 421 selected
Select Reference Samples

Comparison Samples 421 of 421 selected
Select Comparison Samples

Metadata category to color graph by DiseaseState

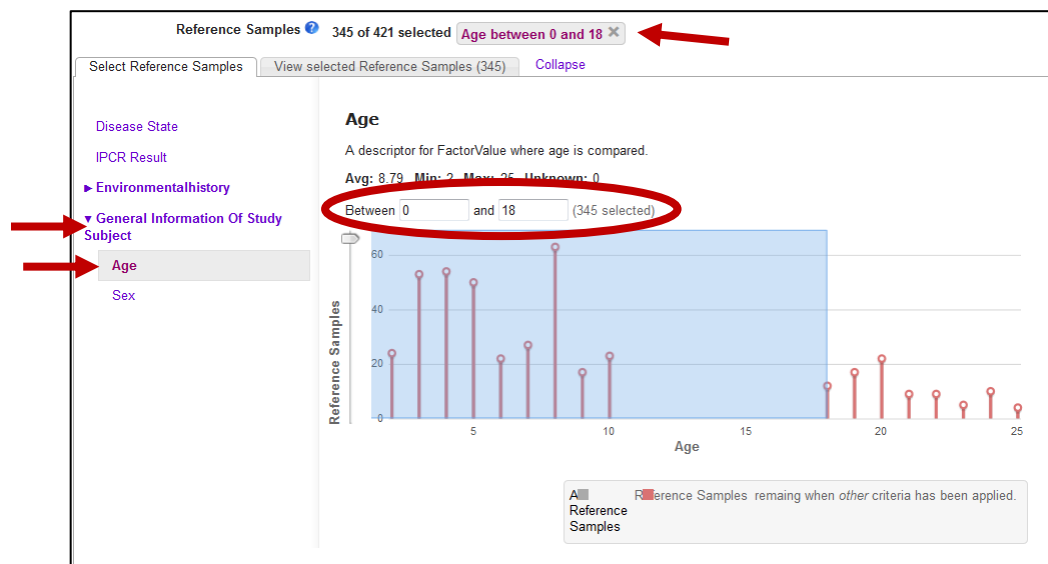
Direction increased immunogenicity

P value less than or equal to 0.05

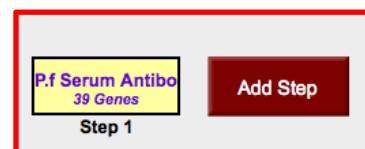
Get Answer

This is a view of the search page with the reference and comparison sample parameters collapsed.

In this example, your **comparison samples will be normal children** and your **reference samples will be infected children**. So each set of samples (reference and comparison) has two parameters that need to be set, age and disease state. The age parameter should be set to 0-18 years for both the reference and the comparison samples. To set the age parameter for the reference samples, choose the 'General Information of Study Subject' and then 'Age' from the left menu to reveal choices for the parameter on the right. Set the range to 0-18 (see image below).



Now set the disease state for the reference samples to infected. Move on the comparison samples and set the age to 0-18 and the disease state to normal. The default settings for other parameters are good – increased immunogenicity and p-value = 0.05.



Reference Samples 265 selected Age is between 0 and 18 Disease State is Infected

Select Reference Samples
expand all | collapse all
Disease State
IPCR Result
Environmental history
General Information Of Study Subject
Age
Sex

View selection (265)
Collapse
Disease State
The name of the pathology diagnosed in the organism from which the biomaterial...
read more
select all | clear all
☒ Infected 272 64.61%
☐ Normal 149 35.39%
All Reference Samples
% Reference Samples from other selected options

Comparison Samples 80 selected Age is between 0 and 18 Disease State is Normal

Refine selection

Metadata category to color graph by DiseaseState

Direction increased immunogenicity

P value less than or equal to 0.05

Advanced Parameters

Get Answer

You are ready to click Get Answer! What do your results look like? Could these represent potential protective antigens? (result image below)

My Strategies:
New
Opened (1)
All (1)
Basket
Public Strategies (33)
Help

(Genes)
Strategy: Serum Ab Response (p-val)
Rename
Duplicate
Save As
Share
Delete

Serum Ab Respo
31 Genes
Step 1
Add Step

31 Genes from Step 1
Strategy: Serum Ab Response (p-val)
Add 31 Genes to Basket | Download 31 Genes

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

Gene Results
Genome View
Analyze Results
BETA

First 1 2 Next Last
Advanced Paging
Add Columns

Gene ID	Organism	Product Description	P Value	Avg Ref ($\text{arcsinh}(1+50x)$)	Avg Comp ($\text{arcsinh}(1+50x)$)	Expression - graph
PF3D7_0304600	<i>P. falciparum</i> 3D7	circumsporozoite (CS) protein (CSP)	1.72E-04	4.696921	5.00801	Signal Intensity Values - $\text{arcsinh}(1+30^{\text{th}} \text{value})$ - PF3D7_0304600 Signal Intensity Samples colored based on DiseaseState

3. Find *falciparum* antigens that may be protective from recurrence of malaria (and potentially re-infection) For this exercise use <http://plasmodb.org>

A recently published study from Kenya ([view paper](#)) followed participants for 12 weeks after an initial screening for malaria and treatment with anti-malarial drugs. The authors collected patient serum, analyzed the serum by protein array and EuPathDB integrated the data set. On a weekly basis, the authors also collected other data associated with the serum sample (metadata) such as 'presence of parasites' and clinical symptoms of malaria. We can use the metadata to stratify the samples and compare subsets of the participants (young vs old, symptomatic vs asymptomatic) when we search for genes that act as antigens. The study concluded:

- increased antigenicity was present in children who did not show clinical symptoms of malaria and suggest that these antigens are protective in children whose malaria symptoms did not recur.
- there was no correlation between antigenicity and time to re-infection

Test both these conclusions.

- a. Compare children (age 0-12) who got clinical malaria during the study (time to first malaria Dx weeks 4-9) to those who didn't (week 11+).
 - Set two parameters for the Reference sample:
Age = 0-12
Time to First Malaria Dx = weeks 4-9.
 - Set two parameters for the Comparison Sample:
Age = 0-12
Time to First Malaria Dx = Week 11+ and Week 11+.
 - Try running with increased immunogenicity then revise and change to decreased immunogenicity. See images below for help configuring the search.
 - Do these results make sense?
- b. Ask the same question (age 0-12) 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.

