Metabolic Pathways
Exploring pathways and compounds

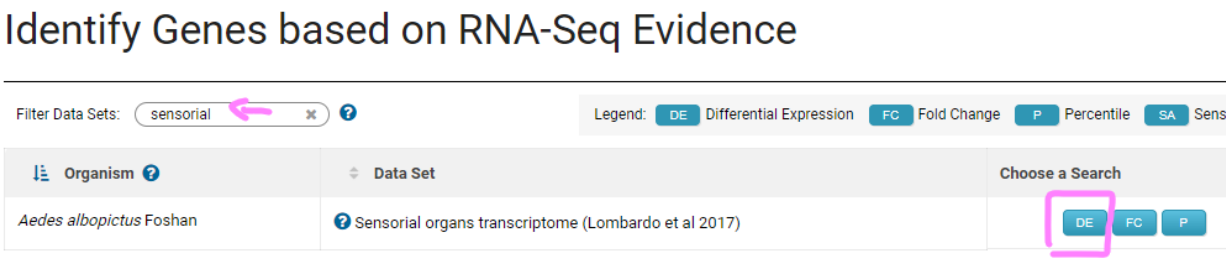
Note: *this exercise uses VectorBase.org and PlasmoDB.org as an example database, but the same functionality is available on all VEuPathDB resources.*

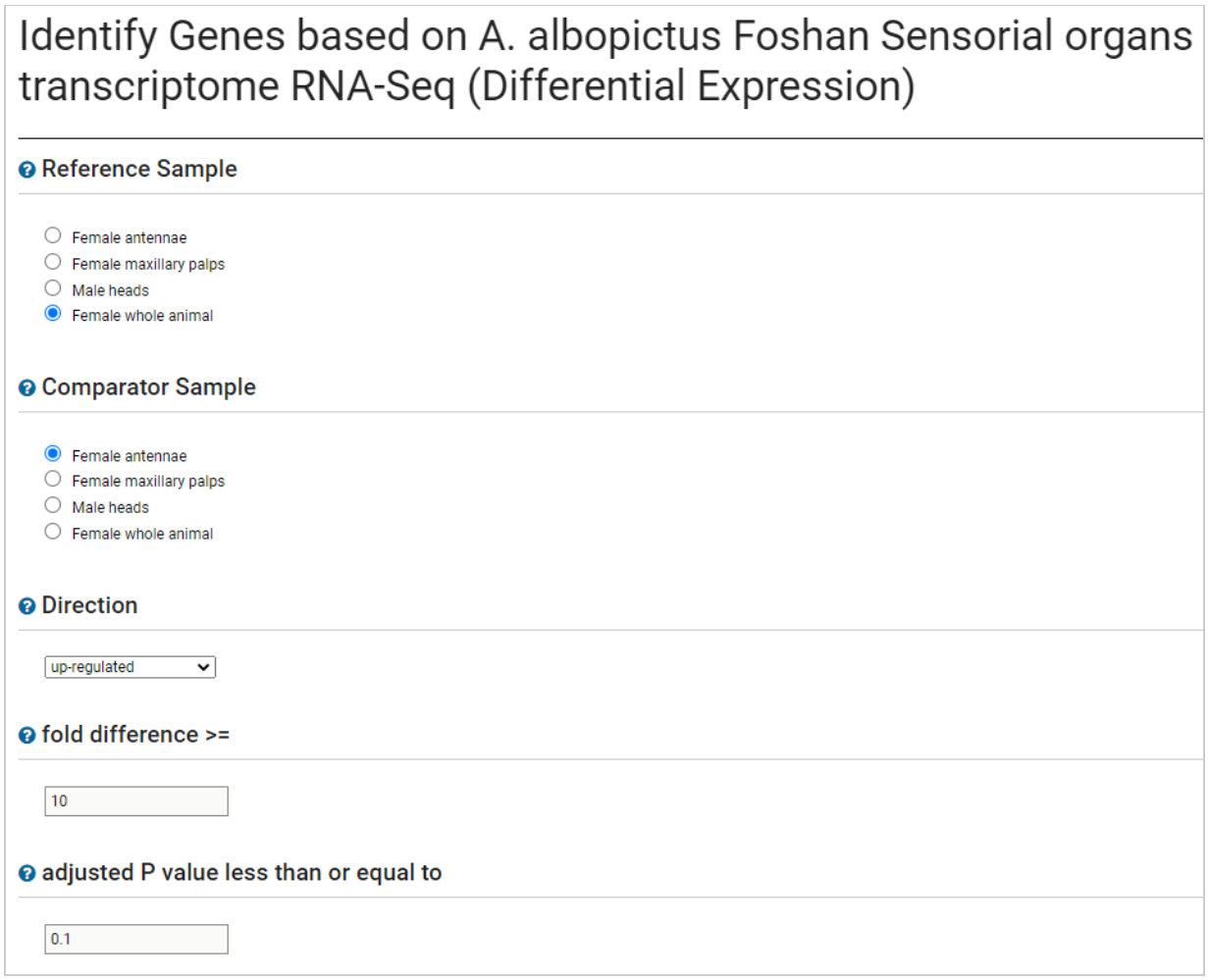
Learning objectives:

- Explore the metabolic pathways searches and visualization tools
- Search for a pathway using the name or pathway identifier
- Paint data onto pathway maps to explore:
    a. Which enzymes in a pathway are present in different experiments
    b. Which enzymes in a pathway are present in different genera
    c. How transcriptional abundance of enzymes in a pathway differs under experimental conditions
- Explore the compound search options

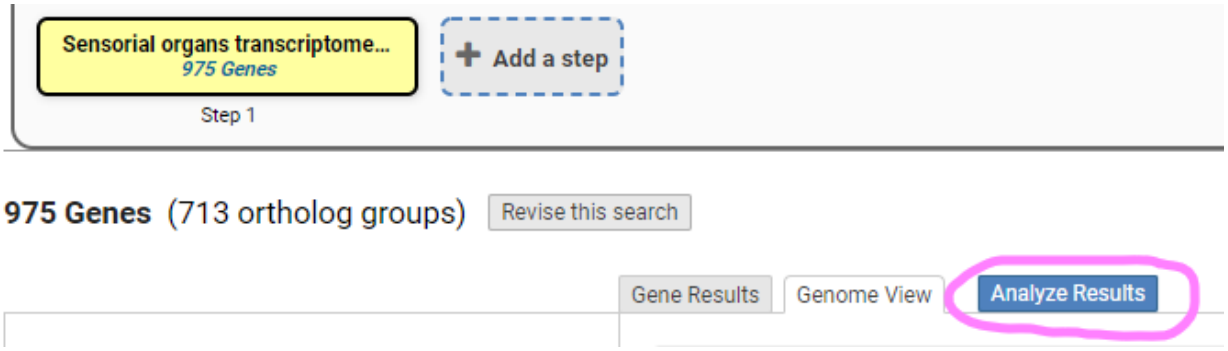
1. **Find genes associated with sensory tissues and the pathways that these represent. For this exercise use http://vectorbase.org**

    - Construct a search strategy with *Aedes albopictus* RNAseq data set "Sensorial organs transcriptome (Lombardo et al 2017)". Using a differential expression query, find genes upregulated at least 10 fold in female antenna when compared to female whole body.

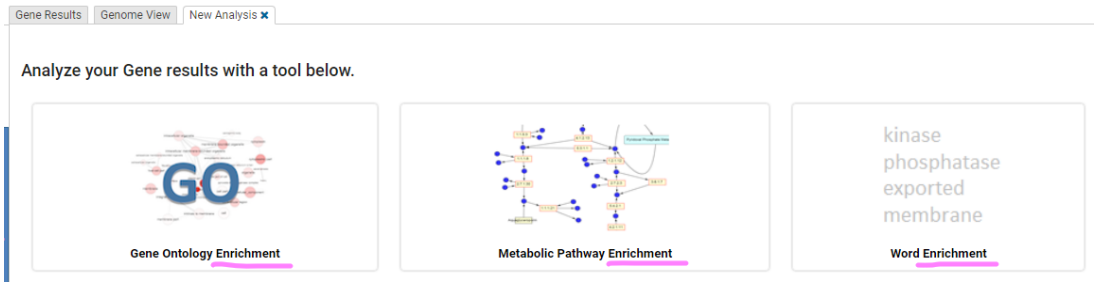




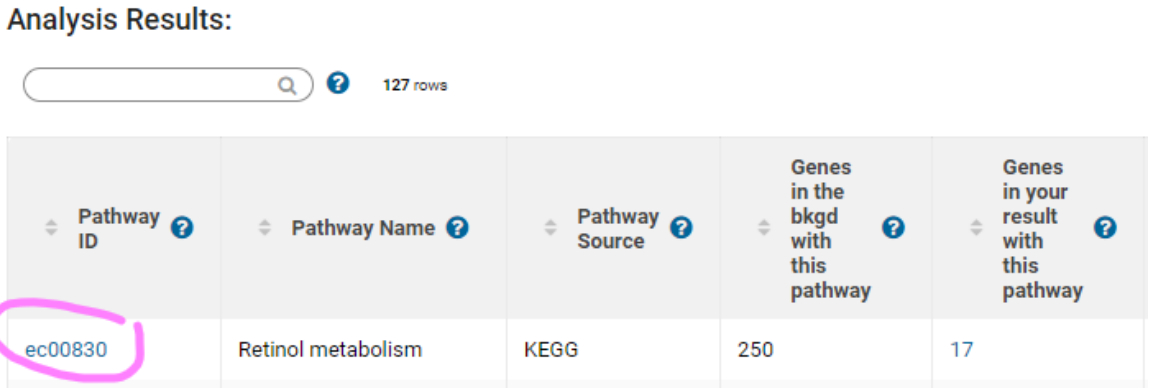
- Click on Analyze Results, to analyse based of three types of enrichment



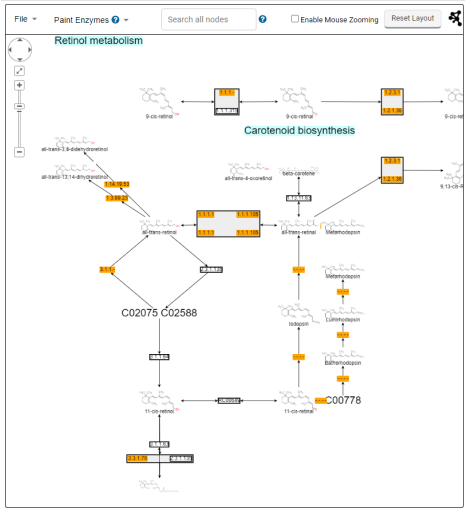
- Word enrichment looks at the text terms associated with the gene product (or function). Give it a try with the default parameters. What is the top hit? Is this an expected result?



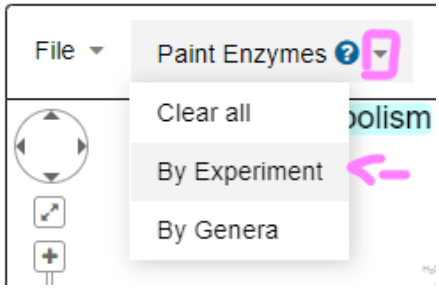
- Run the metabolic pathway enrichment with the defaul parameters. What is the top hit? Is this an expected result? Click on the top hit pathway ID.

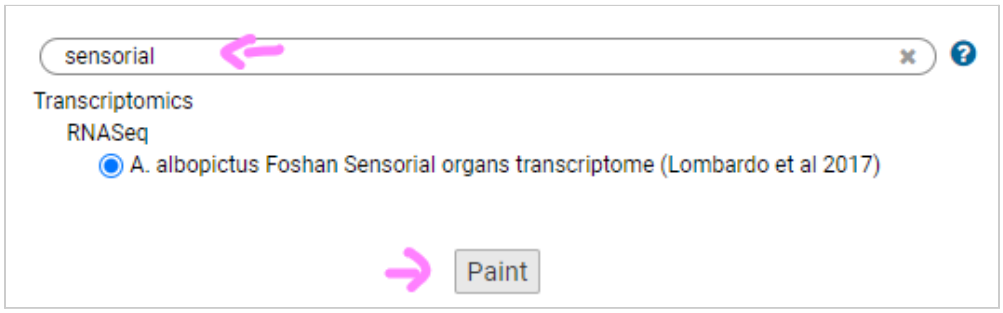


- Scroll down to look at the pathway graphical representation. Enzymes of the pathway present in the *Ae. albopictus* genome are shown in orange color.

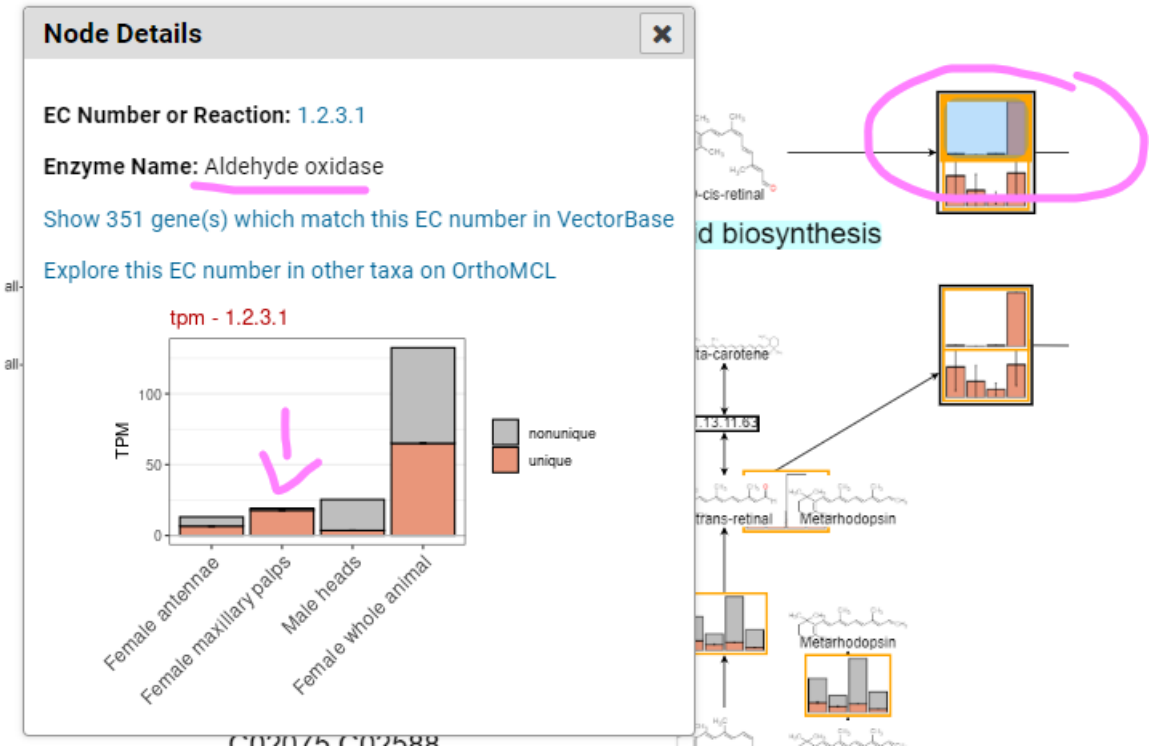


- Paint the enzymes using the experiment used in the search strategy before

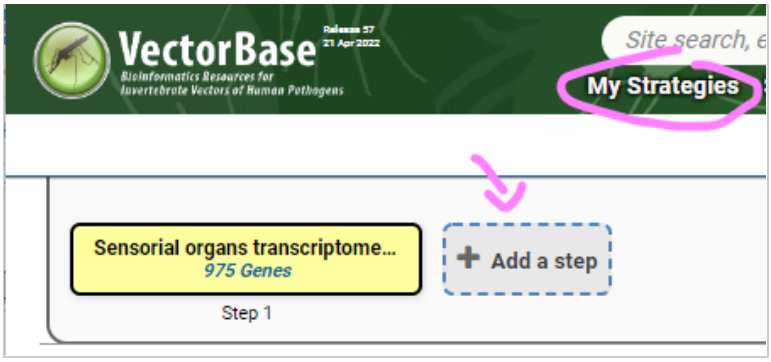




- These graphs might take a little while to load. You can explore the results looking for enzymes patterns of expression in one tissues over the others (excluding the whole body).

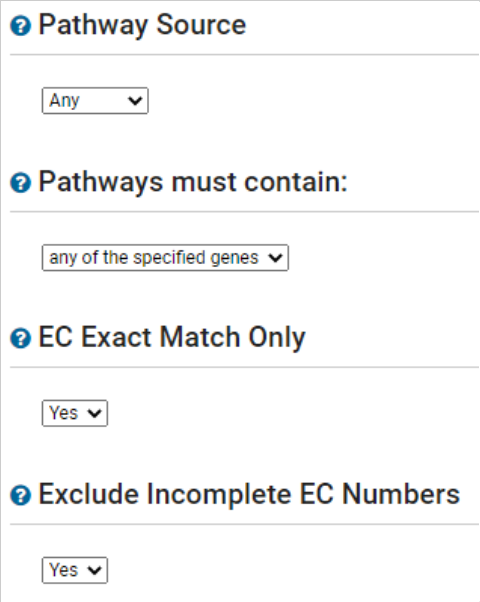


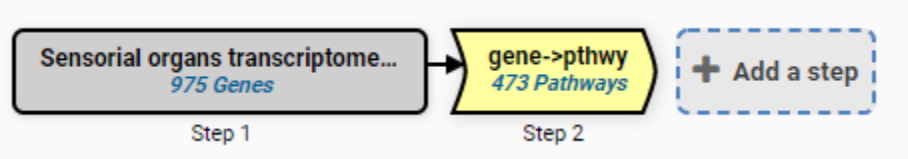
- Now, instead of looking for individual genes and their associated pathways, we can also look at all the pathways associated with the 975 genes. Click on My strategies > Add a step



- Transform the 975 genes into metabolic pathways

- Run the search with the following parameters:



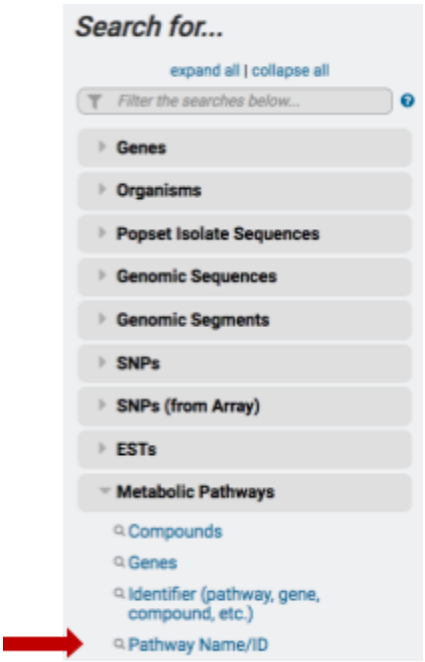


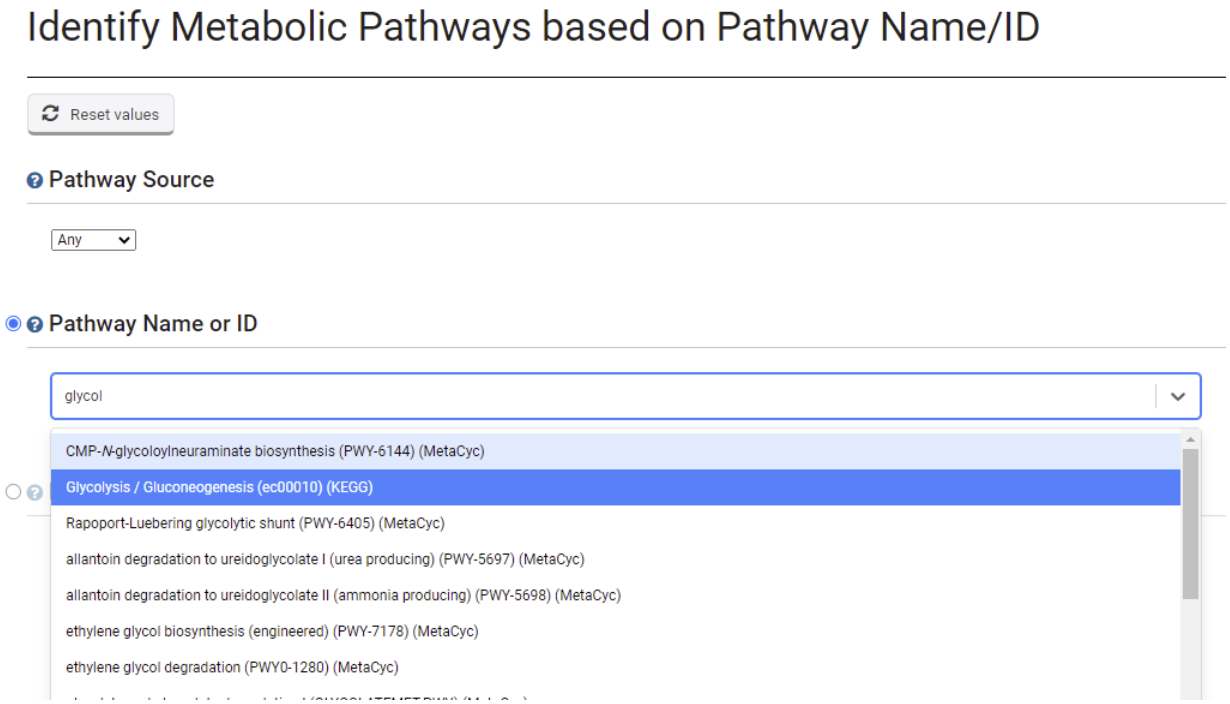
2. Find and explore the metabolic pathway for glycolysis.
For this exercise use http://plasmodb.org

Navigate to the search page for Identify Metabolic Pathways based on Pathway Name/ID.

– Find the metabolic pathway searches on the home page. You can look under "Metabolic Pathways" or use the search filter. You can find metabolic pathways based on the pathway name or identifier, or using genes or compounds involved in the pathway. Search for the **glycolysis** pathway using the Pathway Name/ID option.

– This search is equipped with a type-ahead function for finding the metabolic pathway name. Begin typing glycolysis and then choose the pathway name from the list that appears.

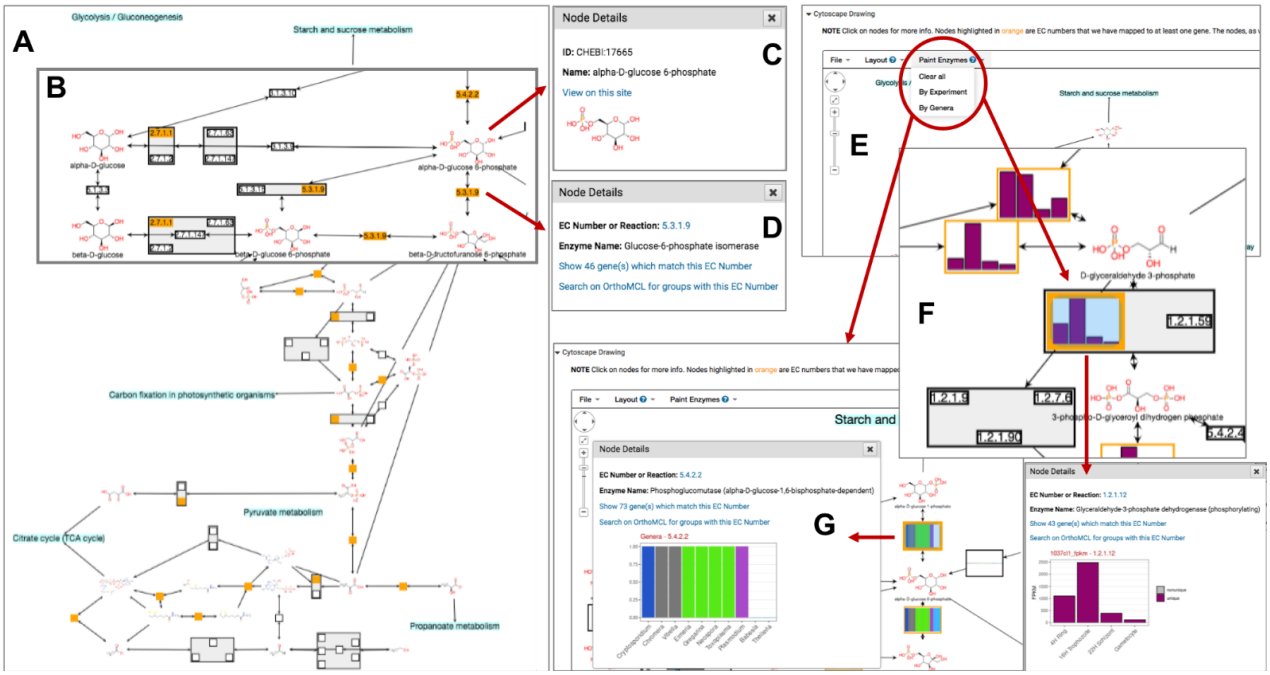


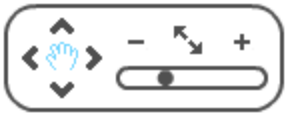


a. Examine the Glycolysis / Gluconeogenesis pathway.

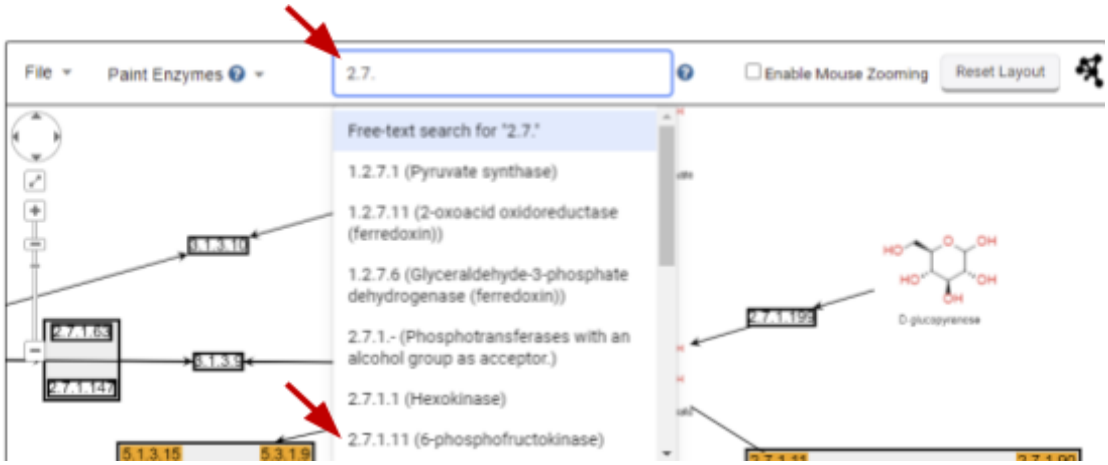
– The search takes you straight to the record page for the Glycolysis / Gluconeogenesis (ec00010) metabolic pathway from KEGG. The Pathways and Interactions section of the record page contains an interactive graphical representation (Cytoscape drawing) of the pathway. The pathway map and the legend can be repositioned.

A. Initial pathway view is zoomed out.
B. Zoom in to see more details including EC numbers and metabolite structures.
C. Click on a compound structure to get additional information.
D. Click on the EC number to get more info about the enzyme including links to retrieve all genes in the database assigned this EC number.
E. The drop-down menu under the heading “Paint Enzymes” allows you paint the pathway based on experimental data or phyletic pattern.
F. Painting the pathway by experiment replaces the enzyme EC numbers with a graphical representation of experimental results for the experiment you choose. Click on the graph to see more details.
G. Painting the pathway based on genera provides a graphical representation of phyletic distribution. Clicking on the phyletic pattern graphic provides additional information.

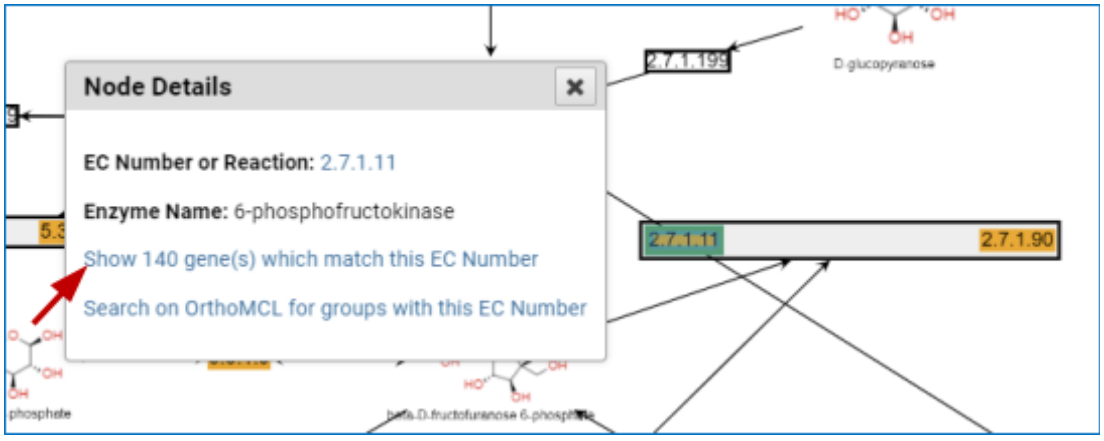




- Use the Tool Box to move (drag) the map and individual nodes. Zoom in and out to help explore the map.
- What do the rectangles with numbers like 2.7.1.11 represent?
- What is the difference between the rectangular nodes that are orange and those that are not?
- Why are some enzymes grouped?
- Find the node representing 6-phosphofructokinase (EC number = 2.7.1.11) using the search in the header of the cytoscape drawing.

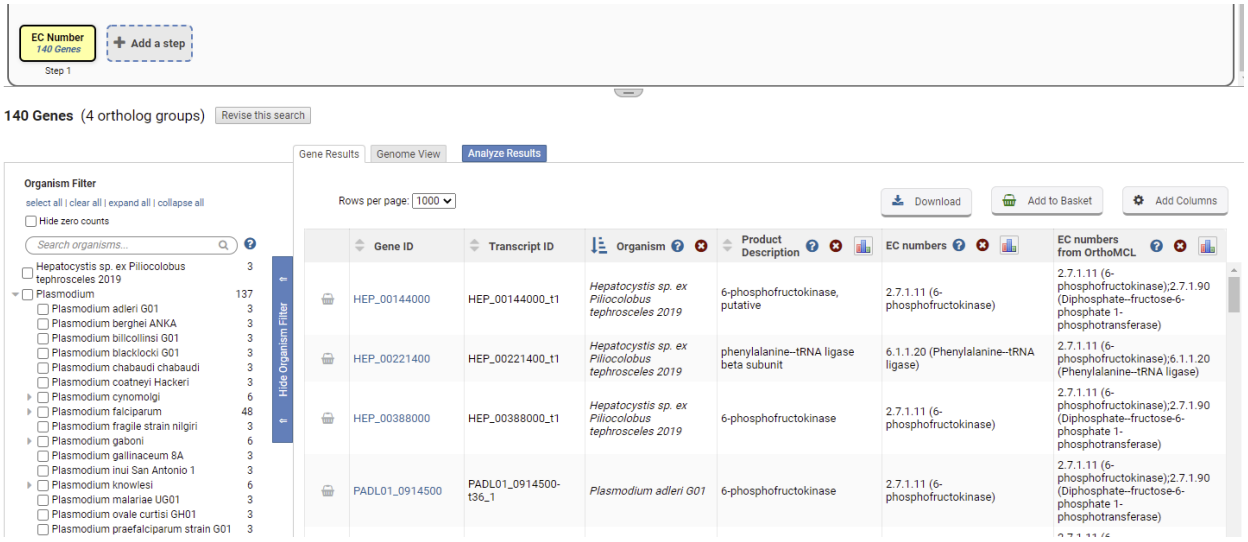


- Click on the 2.7.1.11 node to open a popup with information about this enzyme.



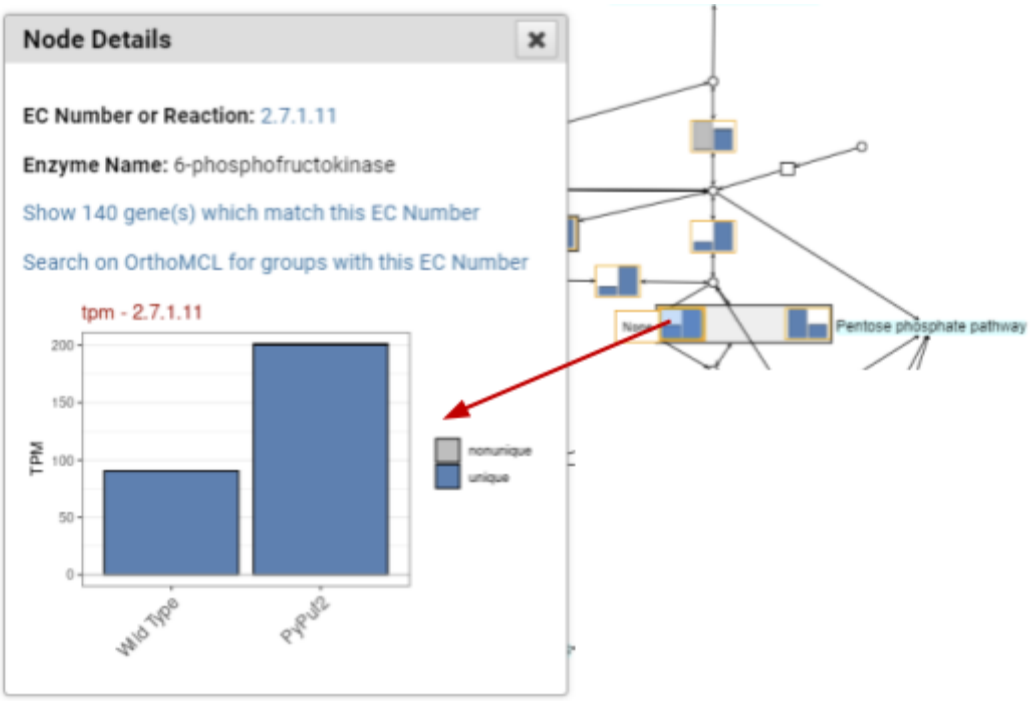
- How many genes in the database matched this EC number?
- Try the link 'Show ### gene(s) which match this EC Number'. Where did you end up? What do the 140 genes in the result list represent? Is 6-phosphofructokinase unique to *P. falciparum*? Notice the two columns called

"EC numbers" and "EC numbers from OrthoMCL". What do these columns represent?

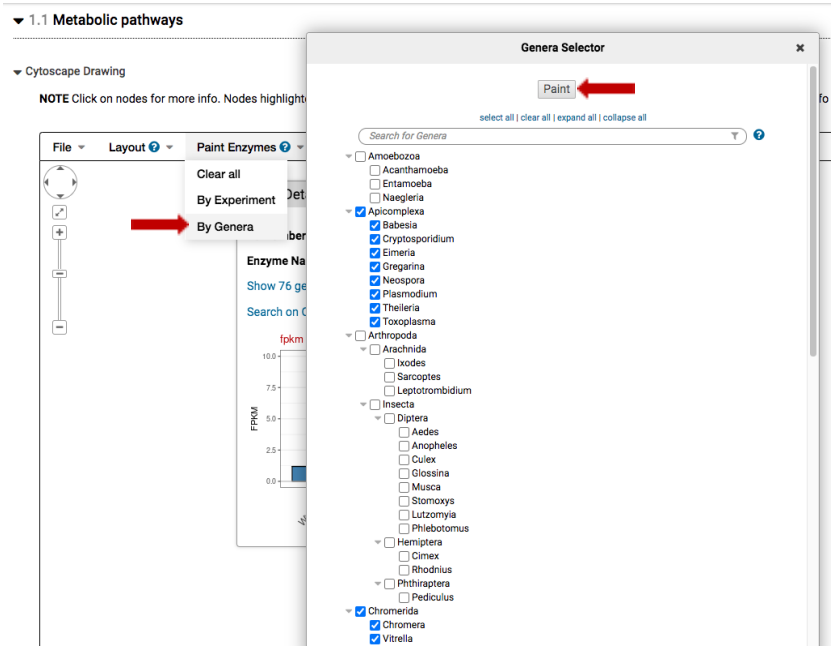


– Use your Browser's back button to return to the glycolysis pathway record page and open the Paint Enzymes menu. Choose 'By Experiment' and select the RNA-seq data set called "Salivary gland sporozoite transcriptomes: WT vs Puf2-KO (Lindner et al)". Be patient while the graphs appear in place of the EC numbers.

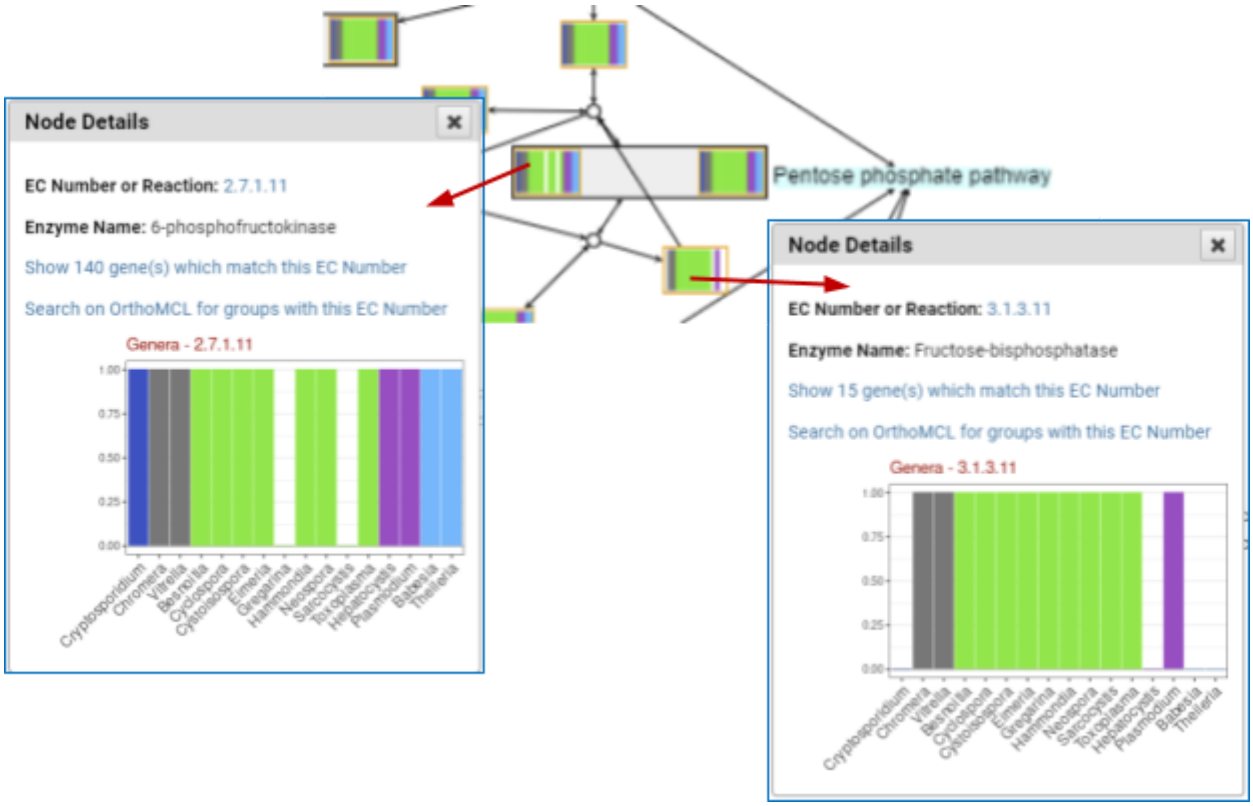
- Does 6-phosphofructokinase appear to be expressed in salivary gland sporozoites? What enzymes in this pathway are affected in knockouts of Puf2?



- Use the Paint Genera option to determine whether 6-phosphofructokinase has orthologs across Apicomplexa and Chromerida.

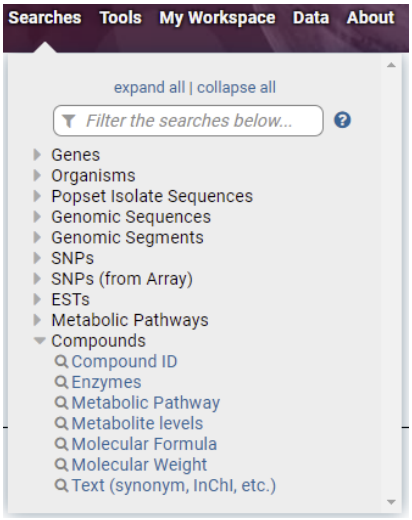


- What about the enzyme that catalyzes the reverse reaction (Fructose-bisphosphatase)?

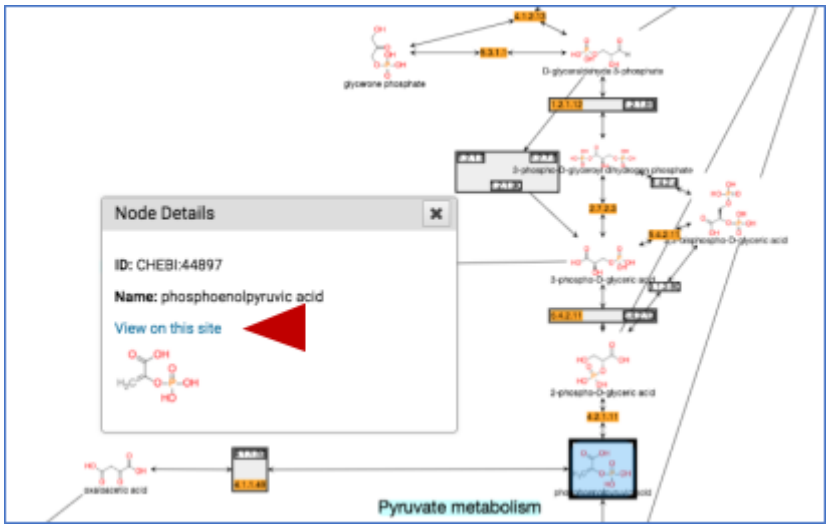


3. **Find and explore the compound record page for phosphoenolpyruvate (phosphoenolpyruvic acid or PEP).**
   Compound records are accessed by running one of the compound searches available under the “Compounds” heading. Compounds may be retrieved by ID, text, metabolic pathway, molecular formula, molecular weight and metabolite levels. Compound records can also be accessed from the metabolic pathway legend after clicking on a compound (blue circle) in the map.

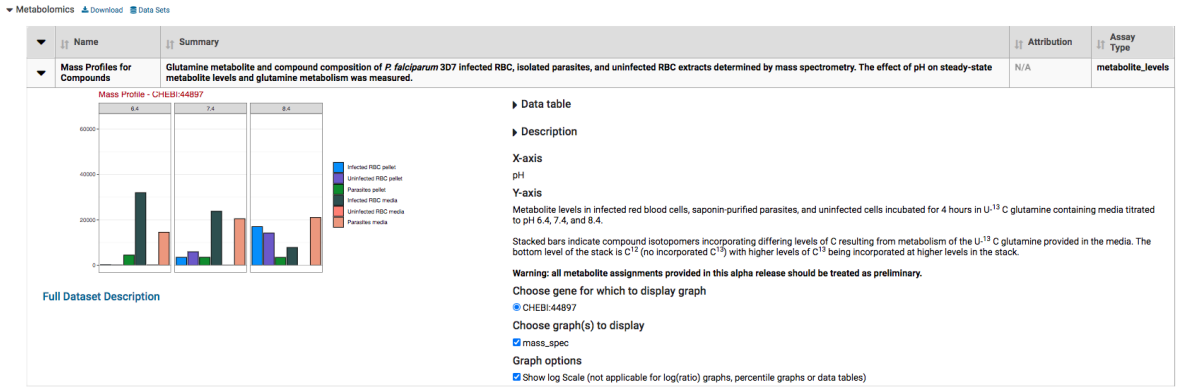
   - Choose one of these searches and retrieve the PEP record page.



- Alternatively, you can reach the PEP record page via a metabolic pathway where it is present as a substrate or a product of an enzymatic reaction (ie. glycolysis). Click on the node representing a compound



- Which method did you use to get to the PEP record page? What compound name worked the best?
- Examine the PEP record page. What data sections do you see?
- Under which conditions is PEP present at highest concentrations? (Hint: navigate to the Metabolomics section)



4. **Find metabolites that are enriched in the isolated parasites (saponin) compared to infected red blood cells (Percoll) that are specific to the cell pellet at pH 7.4**

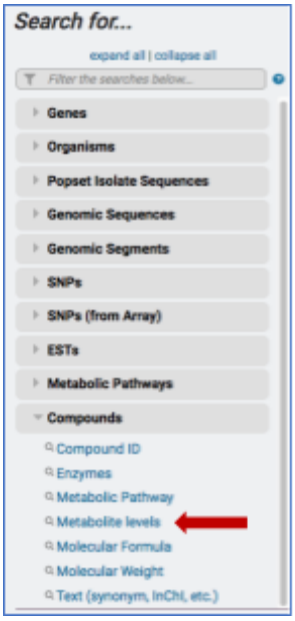
   The metabolite abundance experiment in PlasmoDB compares the 3 conditions at 3 pH levels:
   - Parasites isolated from infected red blood cells using saponin lysis
   - Whole infected red blood cells isolated with Percoll
   - Whole uninfected red blood cells.

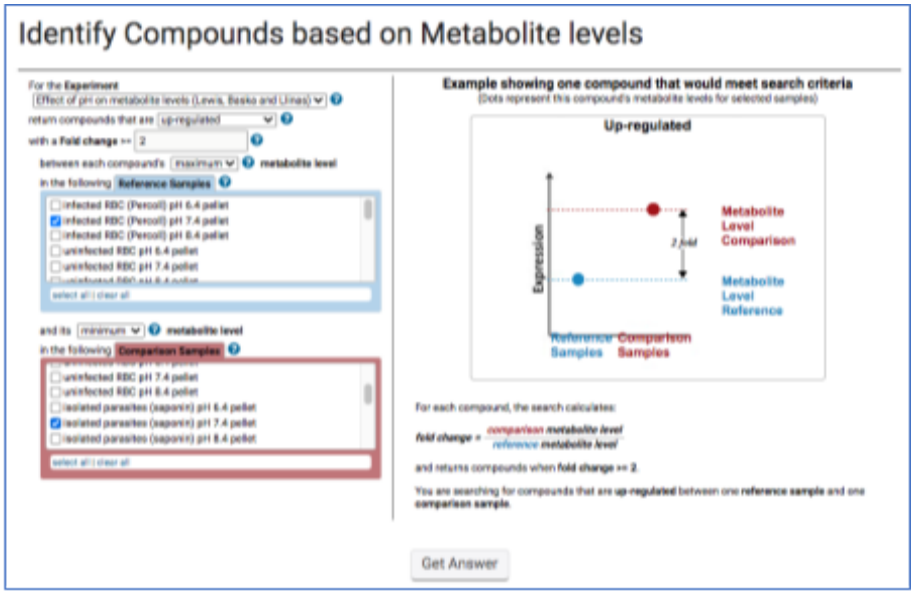
   For all conditions, data was collected from the cell pellet and the media supernatant. Here is the link to the data set record in PlasmoDB https://plasmodb.org/plasmo/app/record/dataset/DS_c3b1287080

   The Metabolite Levels search queries this data set, which uses the same interface as the fold-change searches you have previously seen for transcriptomics data, can be used to find genes whose metabolite levels differ between conditions. Using the strategy system to combine search results it's possible to find genes that are only present in the pellet, by subtracting genes that are also present in the media.

a. Use the Metabolite levels search to find genes that are up-regulated in the pH 7.4 pellets of infected parasite samples compared to infected RBC pH7.4 pellet.
   How many compounds did you get?

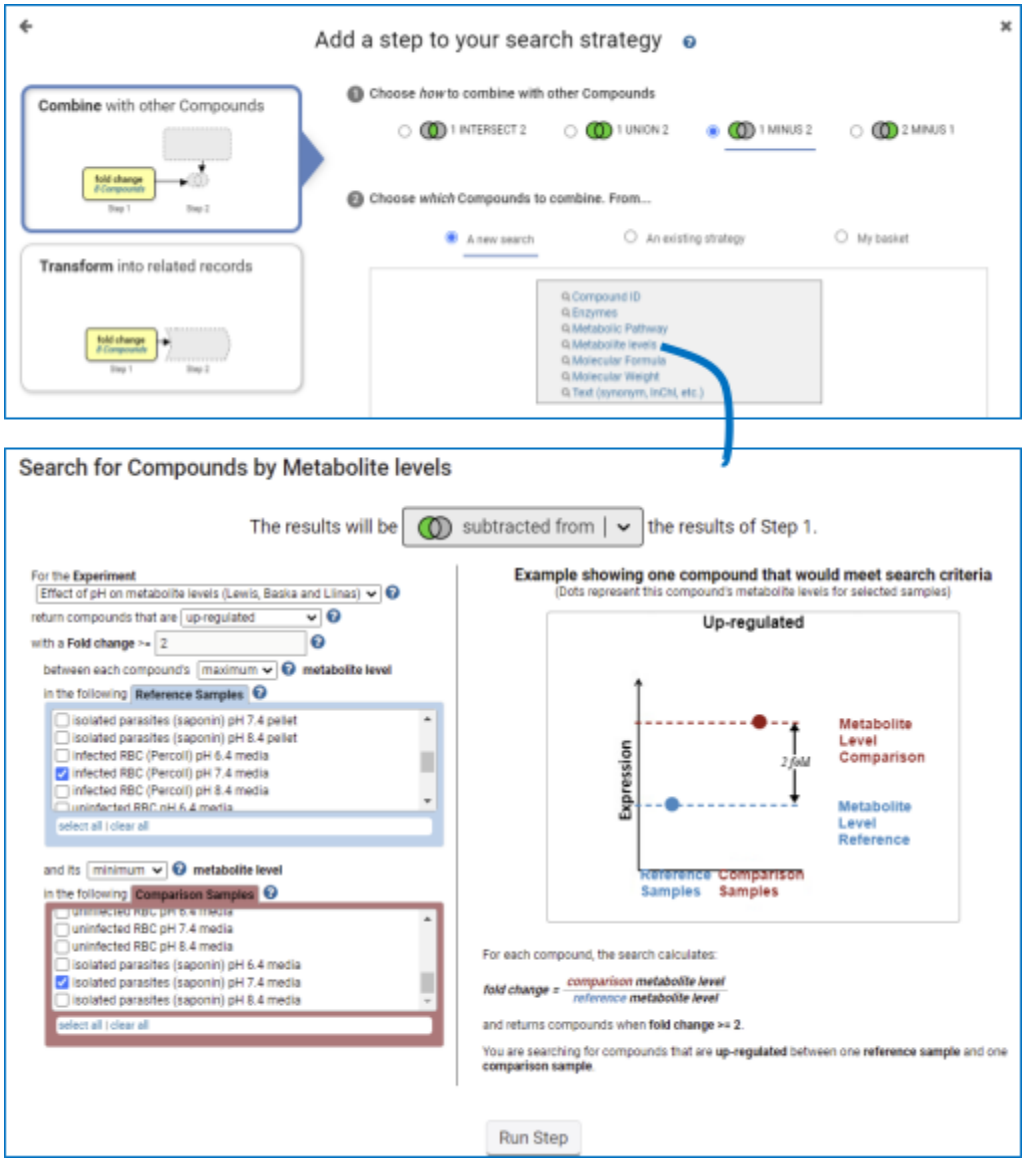
   Fold change >= 2
   (maximum and minimum don't matter here since there is only one sample each)
   Reference = infected RBC pH 7.4 **pellet**
   Comparison = isolated parasites (saponin) pH 7.4 **pellet**





b. Use the strategy system to subtract genes that are also present in the media. Add a step and use the same search to find out how many of these compounds (metabolites) are enriched in the **media supernatant** by 2-fold in isolated parasites (saponin) compared to the infected red blood cells (Percoll).

Fold change >= 2
(maximum and minimum don't matter here since there is only one sample each)
Reference = infected RBC (Percoll) pH 7.4 **media**
Comparison = isolated parasites (saponin) pH 7.4 **media**



How many compounds do you have now? Which metabolic pathways do these compounds belong to? Click Add a Step and transform the results to metabolic pathways.