

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.

Identify Genes based on RNA-Seq Evidence

Filter Data Sets:  

Legend: DE Differential Expression FC Fold Change P Percentile SA Sens

Organism 	Data Set	Choose a Search
<i>Aedes albopictus</i> Foshan	 Sensorial organs transcriptome (Lombardo et al 2017)	<div> DE FC P </div>

Identify Genes based on A. albopictus Foshan Sensorial organs transcriptome RNA-Seq (Differential Expression)

? Reference Sample

- ☐ Female antennae
- ☐ Female maxillary palps
- ☐ Male heads
- ☒ Female whole animal

? Comparator Sample

- ☒ Female antennae
- ☐ Female maxillary palps
- ☐ Male heads
- ☐ Female whole animal

? Direction

up-regulated ▼

? fold difference >=

10

? adjusted P value less than or equal to

0.1

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

Sensorial organs transcriptome...
975 Genes

+ Add a step

Step 1

975 Genes (713 ortholog groups)

Revise this search

Gene Results


Genome View

Analyze Results


- 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?

Gene Results | Genome View | New Analysis ✕

Analyze your Gene results with a tool below.



Gene Ontology Enrichment



Metabolic Pathway Enrichment

kinase
phosphatase
exported
membrane

Word Enrichment

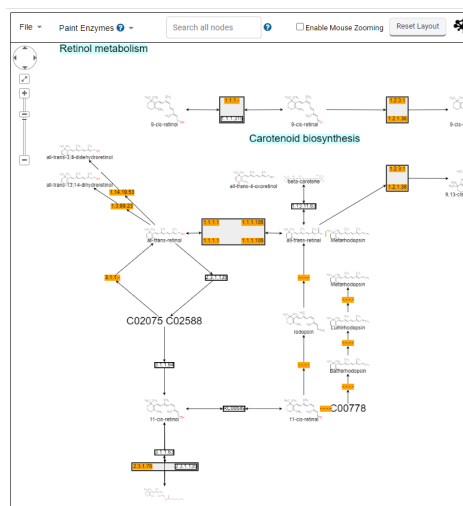
- Run the metabolic pathway enrichment with the default parameters. What is the top hit? Is this an expected result? Click on the top hit pathway ID.

Analysis Results:

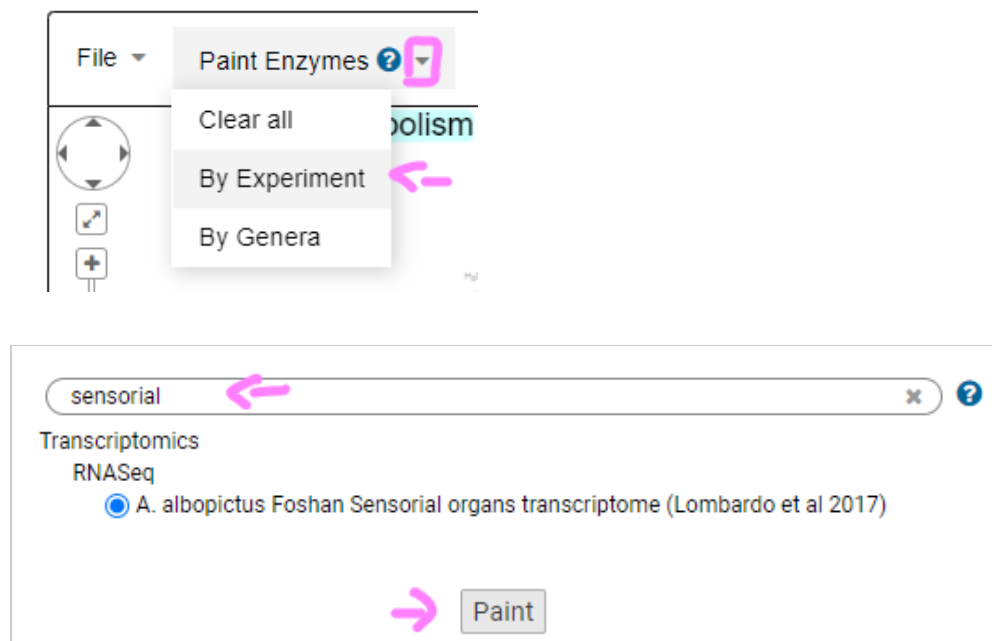
127 rows

Pathway ID ?	Pathway Name ?	Pathway Source ?	Genes in the bkgd with this pathway ?	Genes in your result with this pathway ?
ec00830	Retinol metabolism	KEGG	250	17

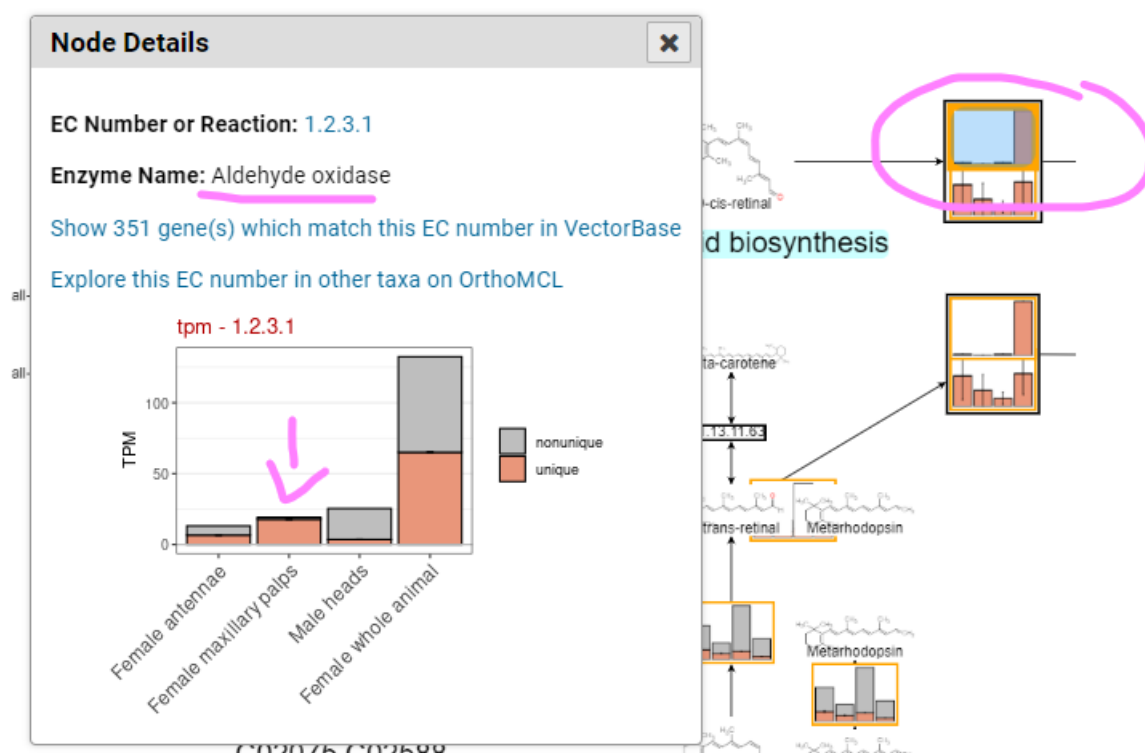
- Scroll down to look at the pathway graphical representation. Enzymes of the pathway present in the *Ae. albopictus* genome are shown in orange color.



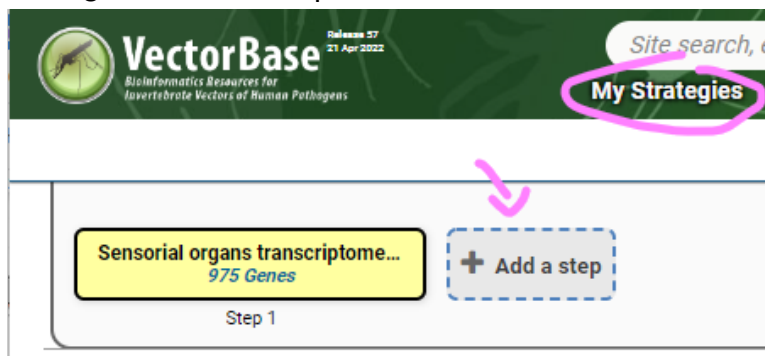
- Paint the enzymes using the experiment used in the search strategy before



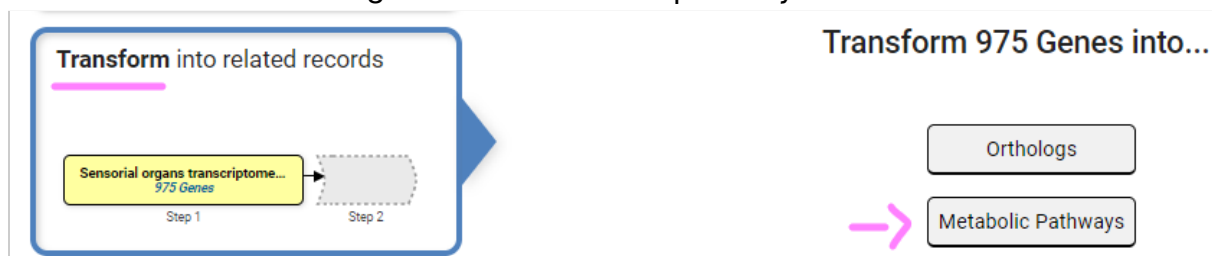
- These graphs might take a little while to load. You can explore the results looking for enzymes patterns of expression in one tissues over the others (excluding the whole body).



- 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:

Pathway Source

Any ▼

Pathways must contain:

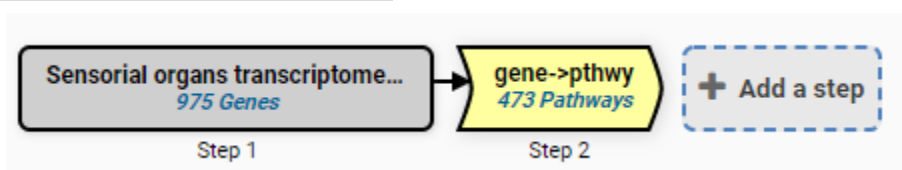
any of the specified genes ▼

EC Exact Match Only

Yes ▼

Exclude Incomplete EC Numbers

Yes ▼



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.

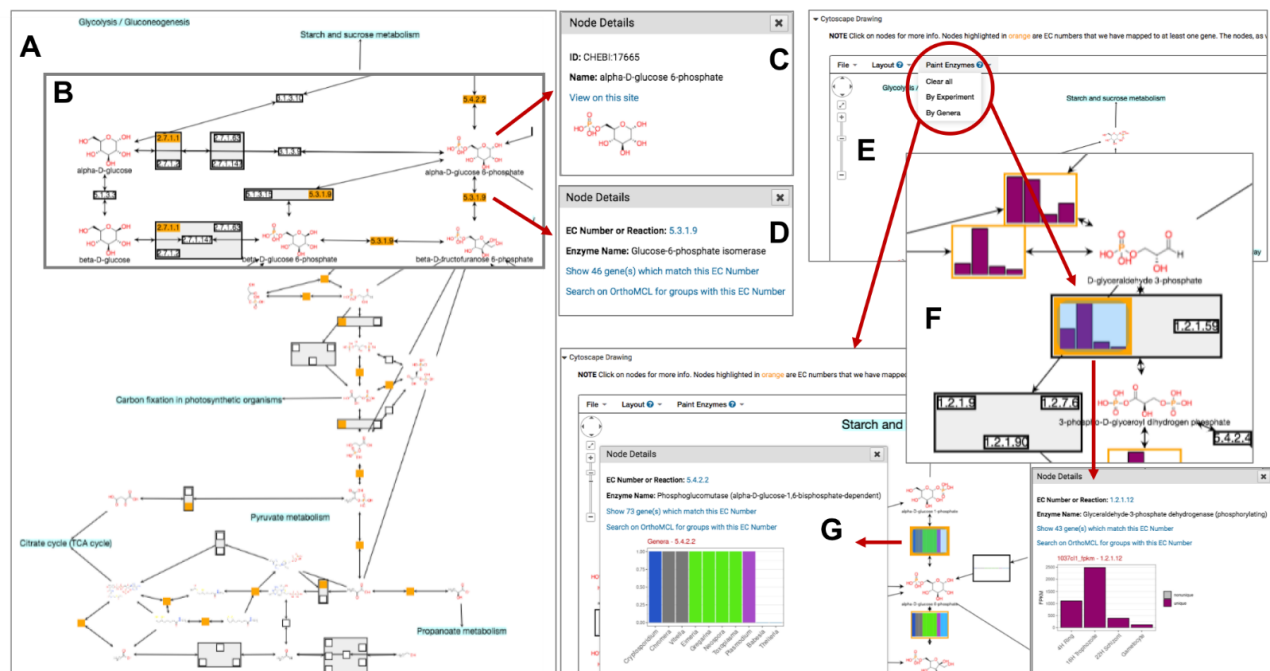
C. Click on a compound structure to get additional information.

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.

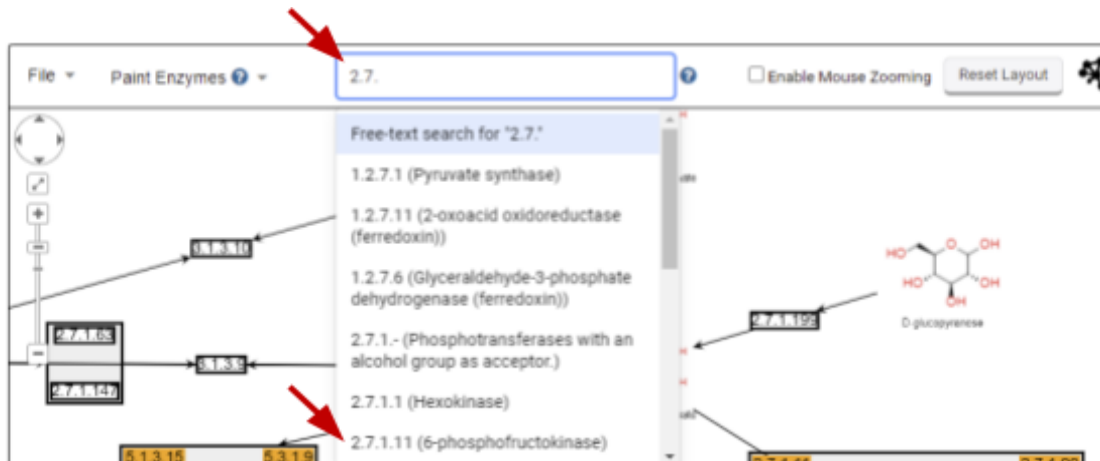
E. The drop-down menu under the heading “Paint Enzymes” allows you paint the pathway based on experimental data or phyletic pattern.

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.

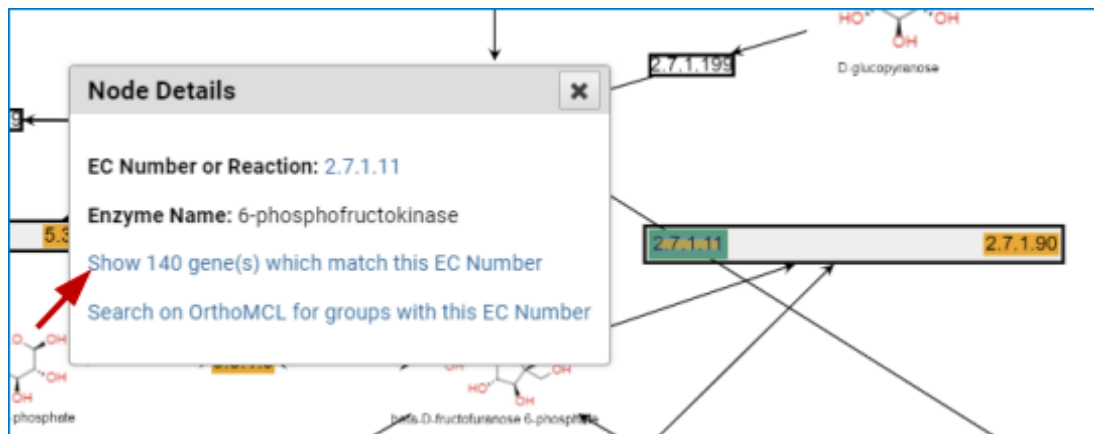
G. Painting the pathway based on genera provides a graphical representation of phyletic distribution. Clicking on the phyletic pattern graphic provides additional information.



- 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?

EC Number
140 Genes
Step 1

+ Add a step

140 Genes (4 ortholog groups) [Revise this search](#)

Gene Results **Genome View** **Analyze Results**

Rows per page: 1000

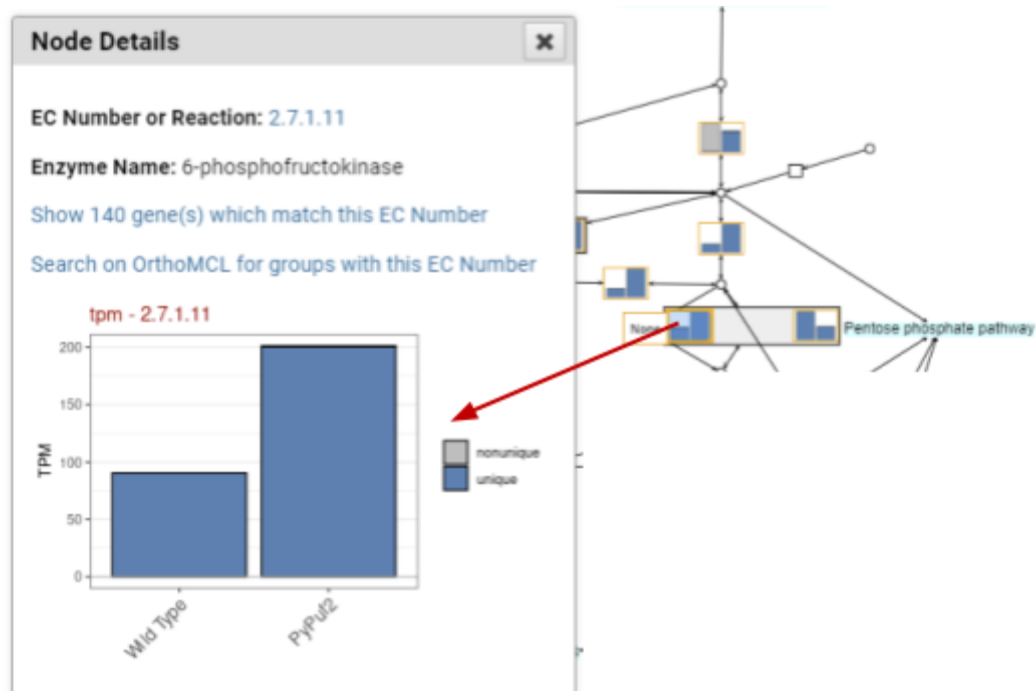
Download Add to Basket Add Columns

Organism Filter
select all | clear all | expand all | collapse all
☐ Hide zero counts
Search organisms...

- ☐ Hepatocystis sp. ex Piliocolobus tephrosceles 2019 3
- ☒ Plasmodium 137
 - ☐ Plasmodium adleri G01 3
 - ☐ Plasmodium berghei ANKA 3
 - ☐ Plasmodium bilcollinsi G01 3
 - ☐ Plasmodium blackioki G01 3
 - ☐ Plasmodium chabaudi chabaudi 3
 - ☐ Plasmodium coatneyi Hackeri 3
 - ☒ Plasmodium cynomolgi 6
 - ☒ Plasmodium falciparum 48
 - ☐ Plasmodium fragile strain niigiri 3
 - ☒ Plasmodium gaboni 6
 - ☐ Plasmodium gallinaceum 8A 3
 - ☐ Plasmodium inui San Antonio 1 3
 - ☒ Plasmodium knowlesi 6
 - ☐ Plasmodium malariae UG01 3
 - ☐ Plasmodium ovale curtisi GH01 3
 - ☐ Plasmodium praefalciparum strain G01 3

Gene ID	Transcript ID	Organism	Product Description	EC numbers	EC numbers from OrthoMCL
HEP_00144000	HEP_00144000.t1	Hepatocystis sp. ex Piliocolobus tephrosceles 2019	6-phosphofructokinase, putative	2.7.1.11 (6-phosphofructokinase)	2.7.1.11 (6-phosphofructokinase); 2.7.1.90 (Diphosphate-fructose-6-phosphate 1-phosphotransferase)
HEP_00221400	HEP_00221400.t1	Hepatocystis sp. ex Piliocolobus tephrosceles 2019	phenylalanine-tRNA ligase beta subunit	6.1.1.20 (Phenylalanine-tRNA ligase)	2.7.1.11 (6-phosphofructokinase); 6.1.1.20 (Phenylalanine-tRNA ligase)
HEP_00388000	HEP_00388000.t1	Hepatocystis sp. ex Piliocolobus tephrosceles 2019	6-phosphofructokinase	2.7.1.11 (6-phosphofructokinase)	2.7.1.11 (6-phosphofructokinase); 2.7.1.90 (Diphosphate-fructose-6-phosphate 1-phosphotransferase)
PADL01_0914500-t36_1	PADL01_0914500-t36_1	Plasmodium adleri G01	6-phosphofructokinase	2.7.1.11 (6-phosphofructokinase)	2.7.1.11 (6-phosphofructokinase); 2.7.1.90 (Diphosphate-fructose-6-phosphate 1-phosphotransferase)

- 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.

1.1 Metabolic pathways

Cytoscape Drawing

NOTE Click on nodes for more info. Nodes highlighted

Paint Enzymes

Clear all

By Experiment

By Genera

Enzyme Name

Show 76 genes

Search on Cytoscape

Genera Selector

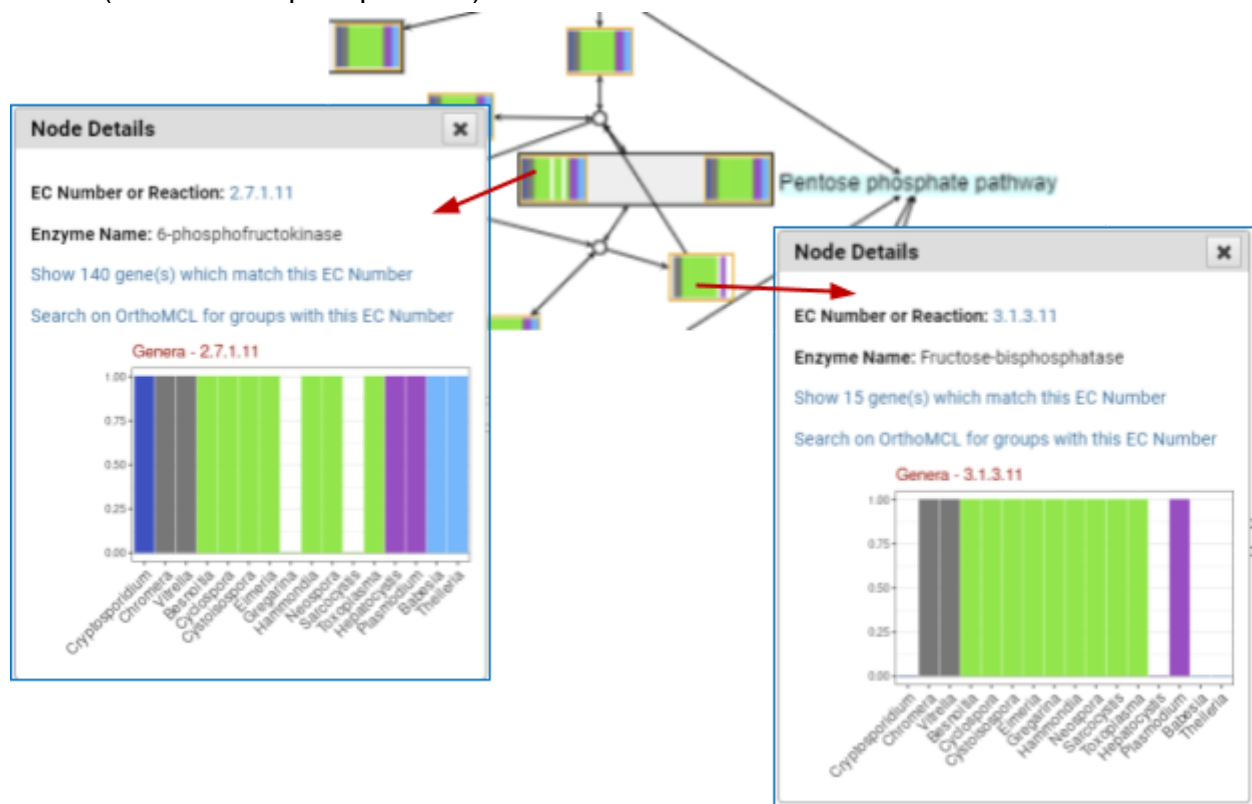
Paint

select all | clear all | expand all | collapse all

Search for Genera

- ☐ Amoebozoa
 - ☐ Acanthamoeba
 - ☐ Entamoeba
 - ☐ Naegleria
- ☒ Apicomplexa
 - ☒ Babesia
 - ☒ Cryptosporidium
 - ☒ Eimeria
 - ☒ Gregarina
 - ☒ Neospora
 - ☒ Plasmodium
 - ☒ Theileria
 - ☒ Toxoplasma
- ☐ Arthropoda
 - ☐ Arachnida
 - ☐ Ixodes
 - ☐ Sarcoptes
 - ☐ Leptotrombidium
 - ☐ Insecta
 - ☐ Diptera
 - ☐ Aedes
 - ☐ Anopheles
 - ☐ Culex
 - ☐ Glossina
 - ☐ Musca
 - ☐ Stomoxys
 - ☐ Lutzomyia
 - ☐ Phlebotomus
 - ☐ Hemiptera
 - ☐ Cimex
 - ☐ Rhodnius
 - ☐ Phthiraptera
 - ☐ Pediculus
- ☒ Chromerida
 - ☒ Chromera
 - ☒ Vitrella

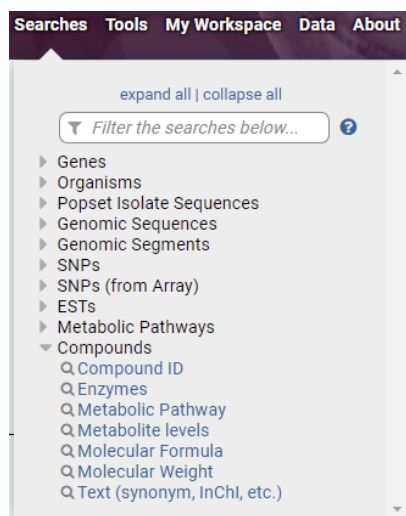
- What about the enzyme that catalyzes the reverse reaction (Fructose-bisphosphatase)?



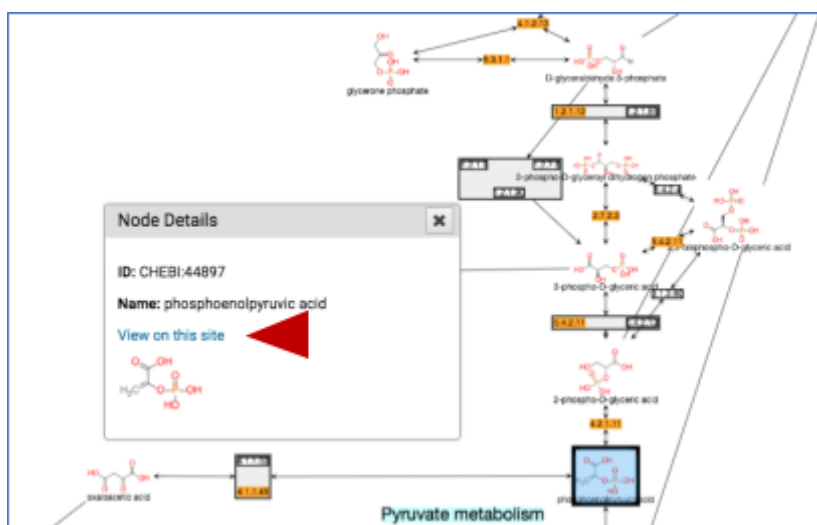
3. Find and explore the compound record page for phosphoenolpyruvate (phosphoenolpyruvic acid or PEP).

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.

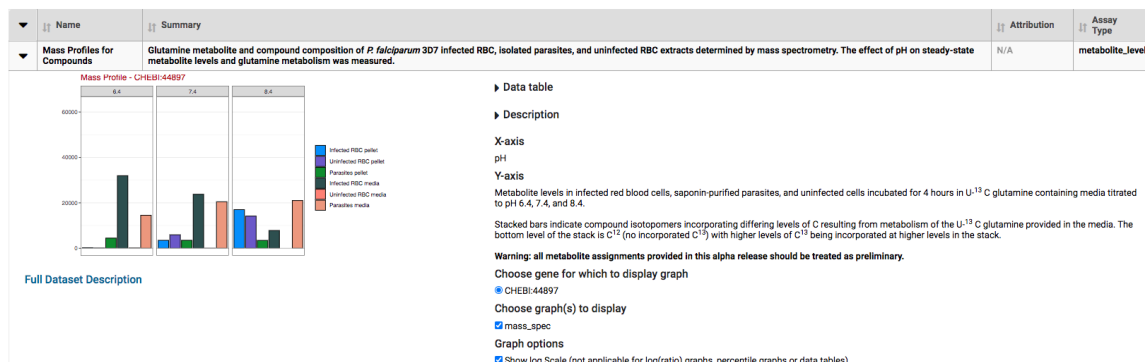


- Alternatively, you can reach the PEP record page via a metabolic pathway where it is present as a substrate or a product of an enzymatic reaction (ie. glycolysis). Click on the node representing a compound



- 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)

Metabolomics Download Data Sets



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.

- 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**

Search for...

expand all | collapse all

Filter the searches below...

- Genes
- Organisms
- Popset Isolate Sequences
- Genomic Sequences
- Genomic Segments
- SNPs
- SNPs (from Array)
- ESTs
- Metabolic Pathways
- Compounds
 - Compound ID
 - Enzymes
 - Metabolic Pathway
 - Metabolite levels 
 - Molecular Formula
 - Molecular Weight
 - Text (synonym, InChI, etc.)

Identify Compounds based on Metabolite levels

For the Experiment
 (Effect of pH on metabolite levels (Lewis, Basko and Linas))

return compounds that are (up-regulated)

with a fold change ≥ 2

between each compound's (maximum) metabolite level

in the following Reference Samples

- ☐ infected RBC (Percoll) pH 6.4 pellet
- ☒ infected RBC (Percoll) pH 7.4 pellet
- ☐ infected RBC (Percoll) pH 8.4 pellet
- ☐ uninfected RBC pH 6.4 pellet
- ☐ uninfected RBC pH 7.4 pellet
- ☐ uninfected RBC pH 8.4 pellet

select all | clear all

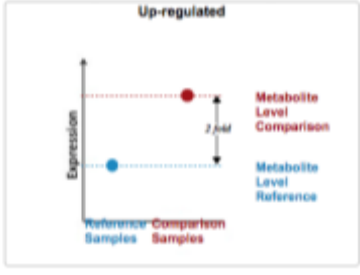
and its (minimum) metabolite level

in the following Comparison Samples

- ☐ uninfected RBC pH 7.4 pellet
- ☐ uninfected RBC pH 8.4 pellet
- ☐ isolated parasites (saponin) pH 6.4 pellet
- ☒ isolated parasites (saponin) pH 7.4 pellet
- ☐ isolated parasites (saponin) pH 8.4 pellet

select all | clear all

Example showing one compound that would meet search criteria
 (dots represent this compound's metabolite levels for selected samples)



For each compound, the search calculates:

$$\text{fold change} = \frac{\text{comparison metabolite level}}{\text{reference metabolite level}}$$
 and returns compounds when fold change ≥ 2 .
 You are searching for compounds that are up-regulated between one reference sample and one comparison sample.

Get Answer

- 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**

Add a step to your search strategy

Combine with other Compounds

1 Choose how to combine with other Compounds

☐ 1 INTERSECT 2
 ☐ 1 UNION 2
 ☒ 1 MINUS 2
 ☐ 2 MINUS 1

2 Choose which Compounds to combine. From...

☒ A new search
 ☐ An existing strategy
 ☐ My basket

Transform into related records

fold change 8 Compounds

fold change 9 Compounds

Search for Compounds by Metabolite levels

The results will be ☒ subtracted from the results of Step 1.

For the Experiment: Effect of pH on metabolite levels (Lewis, Bazila and Llinas)

return compounds that are:

with a Fold change >= 2

between each compound's: metabolite level

In the following Reference Samples:

- ☐ isolated parasites (saponin) pH 7.4 pellet
- ☐ isolated parasites (saponin) pH 8.4 pellet
- ☐ infected RBC (Percoli) pH 6.4 media
- ☒ infected RBC (Percoli) pH 7.4 media
- ☐ infected RBC (Percoli) pH 8.4 media
- ☐ uninfected RBC pH 6.4 media

select all | clear all

and its: metabolite level

In the following Comparison Samples:

- ☐ uninfected RBC pH 6.4 media
- ☐ uninfected RBC pH 7.4 media
- ☐ uninfected RBC pH 8.4 media
- ☐ isolated parasites (saponin) pH 6.4 media
- ☒ isolated parasites (saponin) pH 7.4 media
- ☐ isolated parasites (saponin) pH 8.4 media

select all | clear all

Example showing one compound that would meet search criteria
(Dots represent this compound's metabolite levels for selected samples)

Up-regulated

For each compound, the search calculates:

$$\text{fold change} = \frac{\text{comparison metabolite level}}{\text{reference metabolite level}}$$

and returns compounds when fold change >= 2.

You are searching for compounds that are up-regulated between one reference sample and one comparison sample.

Run Step

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.

