

Metabolic Pathways

Exploring pathways and compounds

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

Learning objectives:

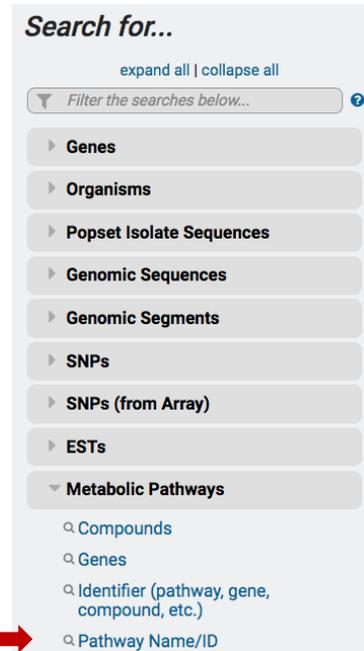
- Explore the metabolic pathways searches and visualization tools
- Search for a pathway using the name or pathway identifier
- Paint data onto pathway maps to explore:
 - a. Which enzymes in a pathway are present in different genera
 - b. How transcriptional abundance of enzymes in a pathway differs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

1. 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.



Identify Metabolic Pathways based on Pathway Name/ID

↻ Reset values

Pathway Source

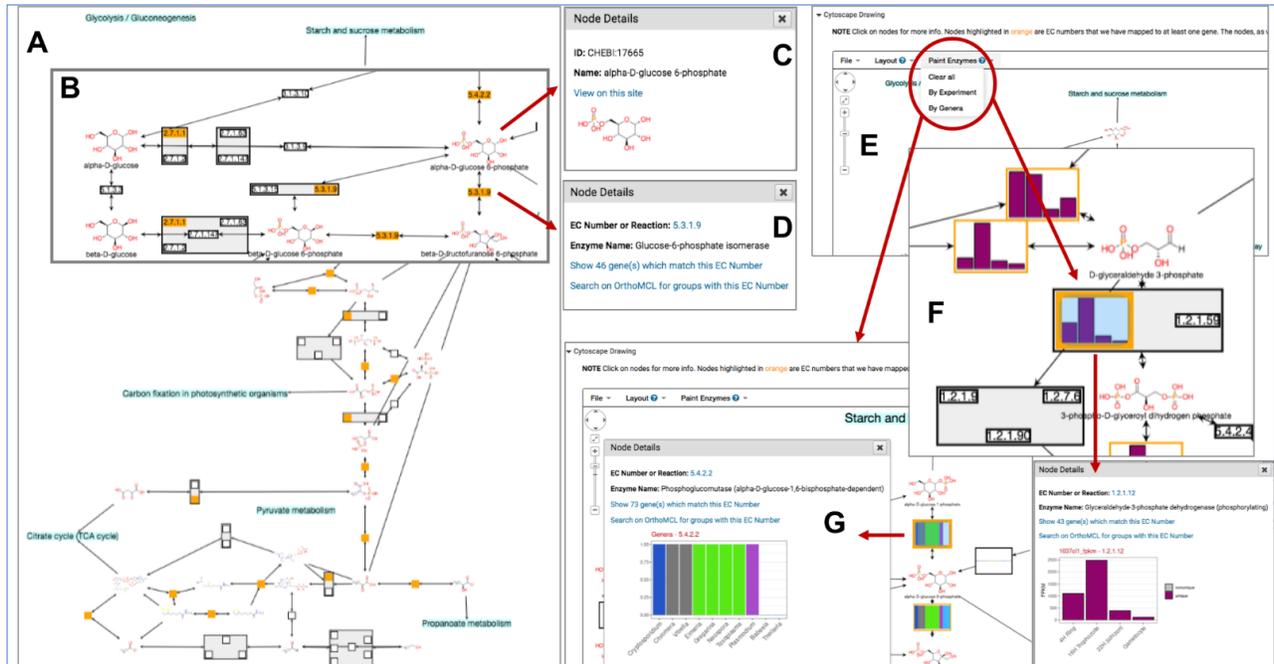
Any ▾

Pathway Name or ID

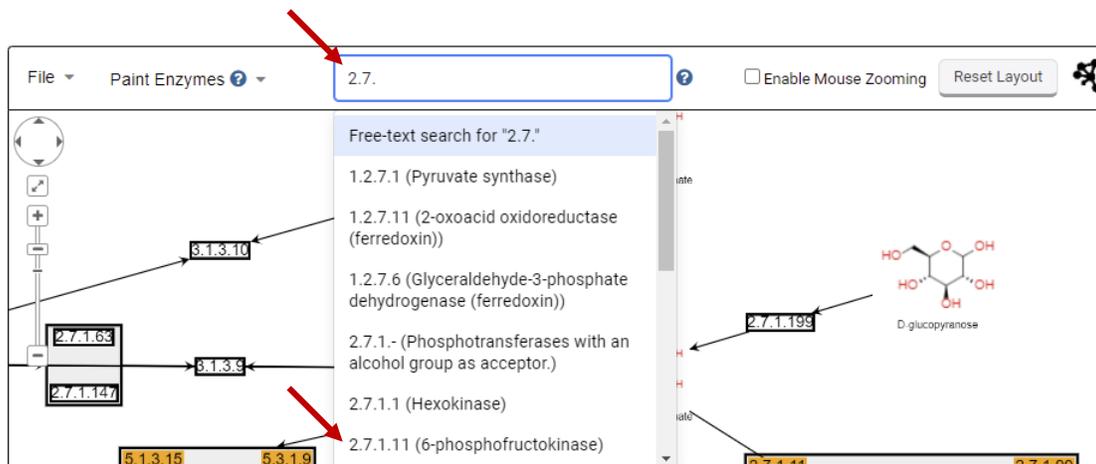
- CMP-N-glycolylneuraminate biosynthesis (PWY-6144) (MetaCyc)
- Glycolysis / Gluconeogenesis (ec00010) (KEGG)**
- Rapoport-Luebering glycolytic shunt (PWY-6405) (MetaCyc)
- allantoin degradation to ureidoglycolate I (urea producing) (PWY-5697) (MetaCyc)
- allantoin degradation to ureidoglycolate II (ammonia producing) (PWY-5698) (MetaCyc)
- ethylene glycol biosynthesis (engineered) (PWY-7178) (MetaCyc)
- ethylene glycol degradation (PWY0-1280) (MetaCyc)

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 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.

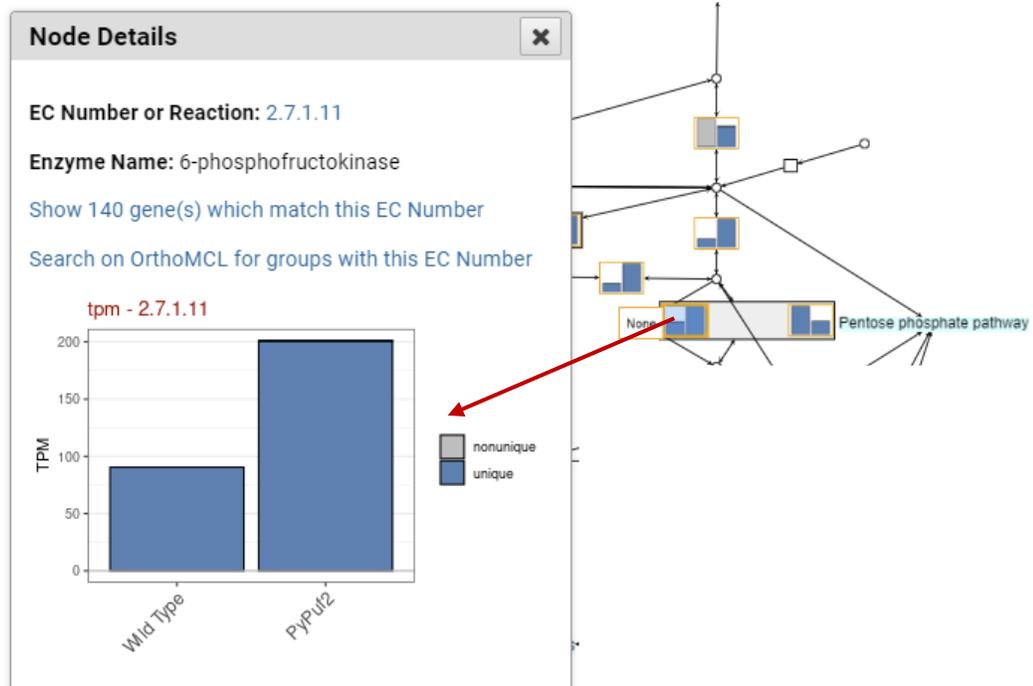


- Click on the 2.7.1.11 node to open a popup with information about this enzyme.

- 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?

Gene ID	Transcript ID	Organism	Product Description	EC numbers	EC numbers from OrthoMCL
HEP_00144000	HEP_00144000_11	<i>Hepatocystis sp. ex Pliococobus tephrosceles 2019</i>	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_11	<i>Hepatocystis sp. ex Pliococobus tephrosceles 2019</i>	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_11	<i>Hepatocystis sp. ex Pliococobus tephrosceles 2019</i>	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	PADL01_0914500-136_1	<i>Plasmodium adleri G01</i>	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

File Layout Paint Enzymes

Clear all

By Experiment

By Genera

Enzyme Name

Show 76 genes

Search on Cytoscape

fpkm

Genera Selector

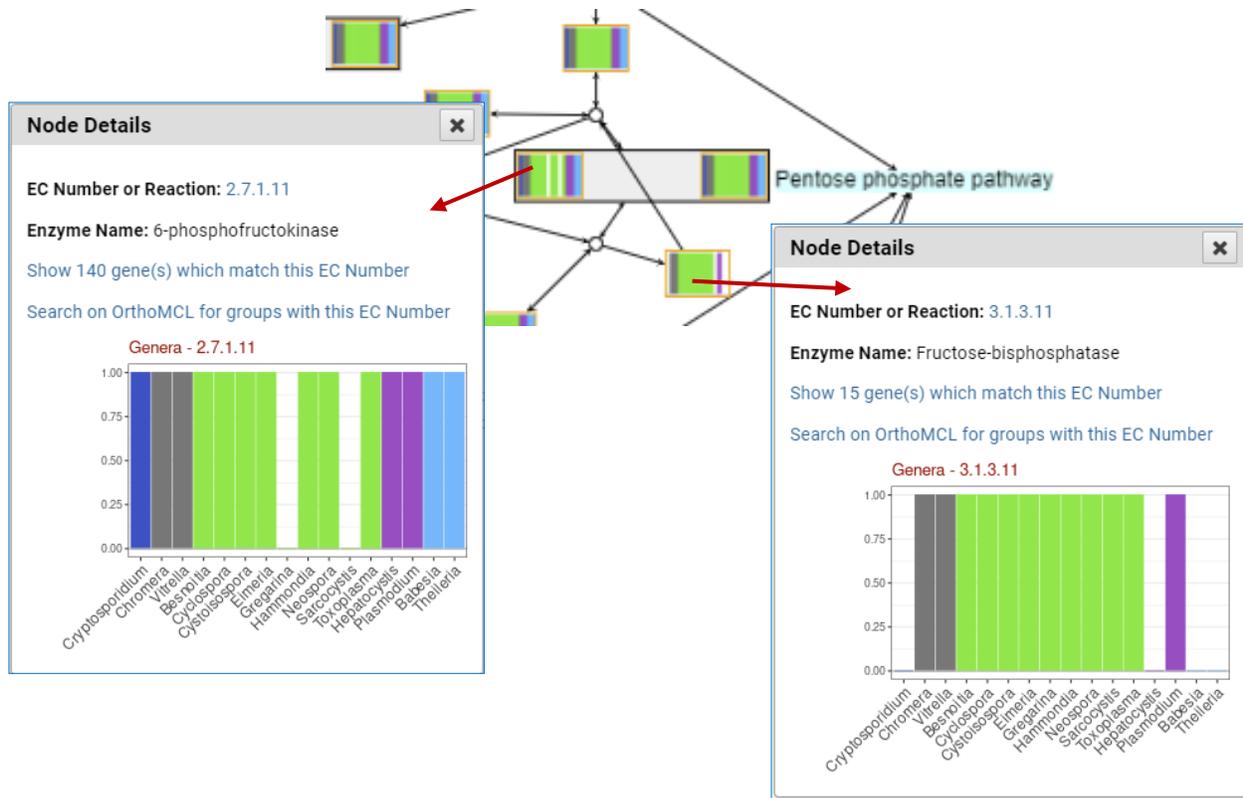
Paint

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
 - Pthiraptera
 - Pediculus
- Chromerida
 - Chromera
 - Vitrella

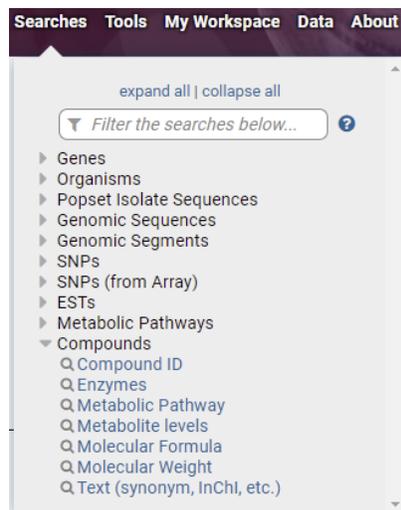
- What about the enzyme that catalyzes the reverse reaction (Fructose-bisphosphatase)?



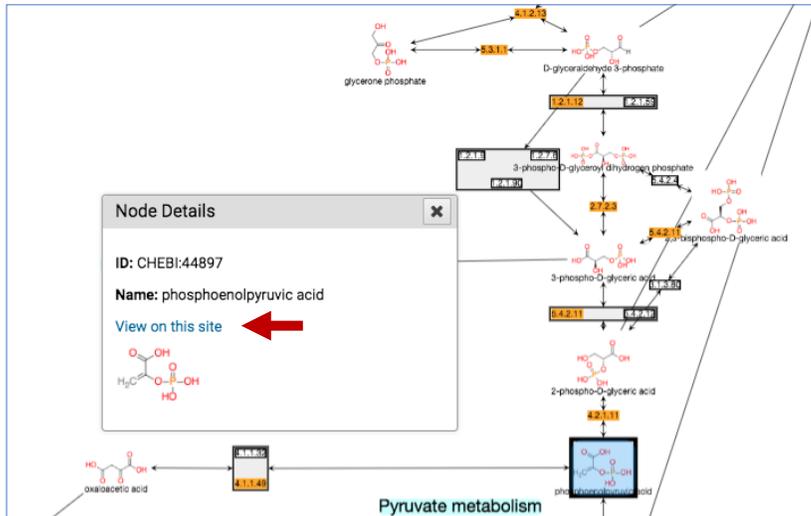
2. Find and explore the compound record page for phosphoenolpyruvate (phosphoenolpyruvic acid or PEP).

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.

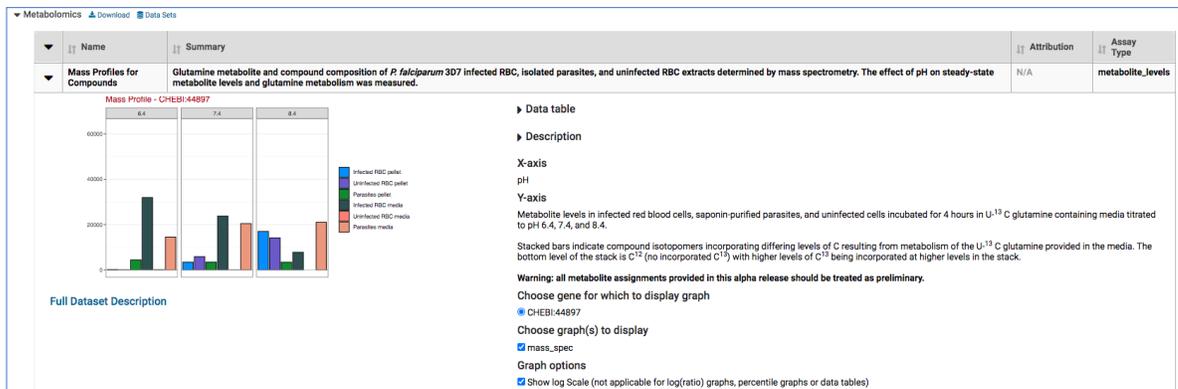
- Choose one of these searches and retrieve the PEP record page.



- 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)



3. 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

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, Baska 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 6.4 pellet
- 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**

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.

