

Mining Proteomics Data

Protein expression in *Leishmania* stages

Learning objectives

- Review the types of expression searches in VEuPathDB
- Use the mass spec evidence search to explore protein expression in metacyclic stages of *Leishmania*

Transcriptomics data in VEuPathDB

Proteomics experiments can determine the peptide sequence by mass spectrometry or the abundance value via e.g. isobaric tagging for a peptide in a sample. Each data type is available on gene record pages but only the mass spec peptides have data that can be viewed in JBrowse. Using the search strategy system, it's easy to delve deep into a specific data set and to take advantage of several types of data when combining search results in the strategy system.

Search	Description	Proteomics
Mass spec. evidence	Peptides from proteomics experiments are mapped to a reference genome enabling searches for genes based on the mapping	✓
Quantitative mass spec. evidence	Quantitative proteomic experiments produce abundance values for proteins identified in the sample analyzed. (excel spreadsheets of gene/protein IDs and expression intensity values)	✓
Post-translational modification (PTM)	PTM data from proteomics experiments, excel spreadsheets of gene/protein IDs, location and type of modification, are associated with genes in VEuPathDB, enabling searches for genes based on the type and number of the PTM.	✓

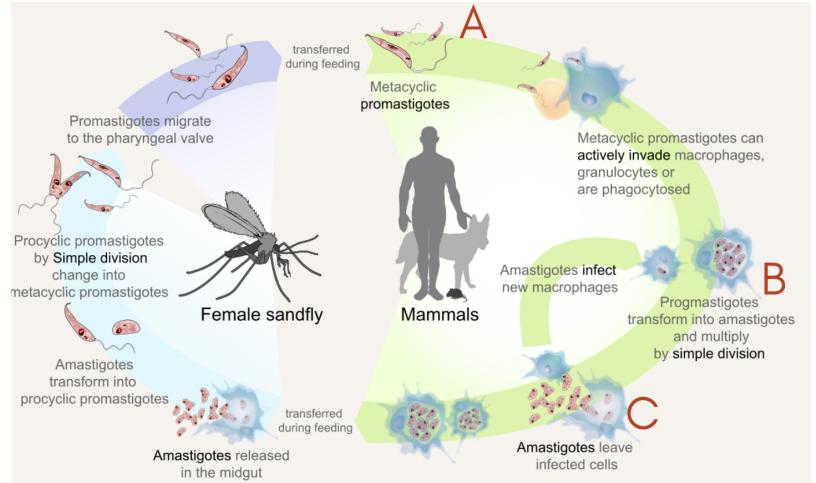
Introduction

Leishmania infantum is the causative agent of infantile visceral leishmaniasis, the most severe form of leishmaniasis, in the Mediterranean region and in Latin America. *L. infantum* has a complex life cycle through sandflies (the vector) and human hosts.

Of the many stages, **metacyclic stage** parasites are the virulent and disease-inducing form of *Leishmania*.

To further characterize this infectious parasitic stage, we could ask the question:

What genes show protein expression in virulent metacyclic stages (A in figure the right) but not in promastigote (B) or amastigote stages (C)?



the
on

1. Navigate to TriTrypDB.org. From the sidebar or header, search or scroll for “proteomics” and click on MassSpec. Evidence

The screenshot shows the TriTrypDB homepage. A search bar at the top contains the word "proteomi". Below it, a dropdown menu shows "proteomi" selected under "Genes". To the right, a sidebar titled "Overview of Resources" lists "Proteomics" under "Genes", with sub-options: "Mass Spec. Evidence", "Post-Translational Modification", and "Quantitative Mass Spec. Evidence". Red arrows point from the search bar to the dropdown menu and from the sidebar to the "Proteomics" link.

2. This opens the search to identify genes based on mass spec. evidence. Filter the experiment and sample tree by typing a keyword in the filter box that represents our species of interest: “infantum”.
 - a. The first step in our strategy is to identify genes that show protein expression in the metacyclic stages.
 - b. Select all *L. infantum* samples that come from the **metacyclic stage** proteome.
 - c. Keep the default search parameters and click on the Get Answer button.

How many genes did you get?

Identify Genes based on Mass Spec. Evidence

Configure Search Learn More View Data Sets Used

Reset values to default

Experiments and Samples

1 selected, out of 152
select only these | add these | clear these | Reference only

Leishmania

Leishmania infantum, JPCM5 (Ouellette, et al. unpublished)

Metacyclic Stage Proteins (Ouellette, et al. unpublished)

Post-translational modified proteins during differentiation (Rosenzweig et al.)

acetylated proteins (L. donovani)

glycosylated proteins (L. donovani)

methylated proteins (L. donovani)

phosphorylated proteins (L. donovani)

Promastigote and Amastigote phosphoproteomes (donovani) (Tsiganek et al.)

amastigote phosphopeptides

promastigote phosphopeptides

Promastigote and amastigote proteomes (MHOM/MA/67/ITMAP-263) (Brotherton et al.)

amastigote by 1DE, LC-MS/MS

amastigote by 2DE, LC-MS/MS, pH6-11

amastigote by 2DE, LC-MS/MS, pH6-9

promastigote by 2DE, LC-MS/MS, pH6-11

promastigote by 2DE, LC-MS/MS, pH6-9

promastigote by 2DE, LC-MS/MS, temp and pH control

promastigote by 2DE, LC-MS/MS, temp and pH stressed

promastigote secretome

Minimum Number of Unique Peptide Sequences

1

Apply min # peptide sequences / sample OR across samples

Per Sample

Advanced Parameters

Get Answer

L. infantum metacyclic protein expression *

Mass Spec 162 Genes

+ Add a step

Step 1

3. We have a set of genes that are expressed in metacyclic stages. To identify genes that are expressed in this stage but not expressed in other stages, we can add a step to the search strategy.

- a. Specifically, subtract the genes that have protein expression in the promastigote and amastigote stages.
- b. To do this, add a step to your strategy that combines the first step with other genes using a 1 minus 2 operator.
- c. To choose which genes to combine, click on Mass Spec. Evidence.

L. infantum metacyclic protein expression *

Mass Spec 162 Genes

+ Add a step

Step 1

Add a step to your search strategy

Choose how to combine with other Genes

- 1 INTERSECT 2
- 1 UNION 2
- 1 MINUS 2
- 2 MINUS 1

Choose which Genes to combine. From...

- A new search
- An existing strategy
- My basket

protein

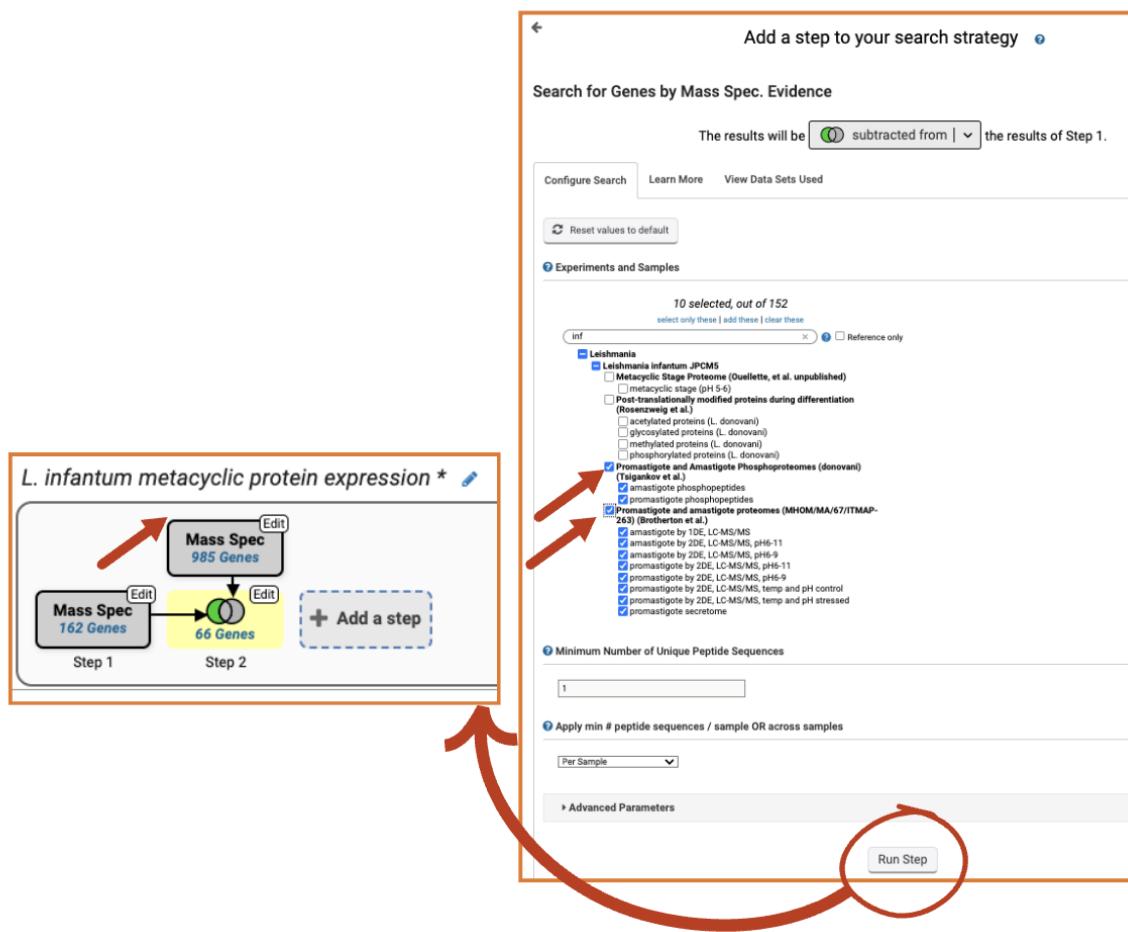
Proteomics

Q: Mass Spec. Evidence

Q: Post-Translational Modification

Q: Quantitative Mass Spec. Evidence

4. Choose all the *L. infantum* samples labeled promastigote and amastigote and run the search.
 How many genes did you get that have evidence of protein expression in metacyclic stages but not in promastigote and amastigote stages?



5. One way to increase stringency is to increase the minimum number of unique peptides that are required to map to a gene before it is returned by the search. The default settings that we used above return any gene with a minimum of 1 peptide.
- Click on the edit button in the first step of the search
 - Click on the Revise option in the popup
 - Change the value of the “Minimum Number of Unique Peptide Sequences” search parameter from 1 to 5 and click on the revise button
 - Remember to do this for each step!

How did this change your results? Would you consider these results more stringent?

L. infantum metacyclic protein expression *

Mass Spec 985 Genes

Mass Spec 162 Genes

66 Genes

Step 1 Step 2

View | Analyze | Revise | Insert step before | Orthologs | Delete

Details for step Mass Spec

162 Genes

Experiments and Samples metacyclic stage (pH 5-6)

Minimum Number of Unique Peptide Sequences 1

sample OR across samples Per Sample

gene (applied per sample) 1

Identify Genes based on Mass Spec. Evidence

Configure Search Learn More View Data Sets Used

Reset values to default

Experiments and Samples

1 selected, out of 152

Leishmania

- Leishmania braziliensis MHOM/BR/75/M2904
- Leishmania donovani BPK28A1
- Leishmania infantum JPCM5
 - Metacyclic Stage Proteome (Duellette, et al. unpublished)
 - Post-translationally modified proteins during differentiation (Rosenzweig et al.)
 - Promastigote and Amastigote Phosphoproteomes (donovani) (Tsiganakis et al.)
 - Promastigote and amastigote proteomes (MHOM/MA/67/ITMAP-263) (Brooker et al.)
 - Leishmania major strain Friedlin
 - Leishmania mexicana MHOM/GT/2001/U1103
 - Trypanosoma

Minimum Number of Unique Peptide Sequences

5

Apply min # peptide sequences / sample OR across samples

Per Sample

Advanced Parameters

Revise

L. infantum metacyclic protein expression *

Mass Spec 248 Genes

Mass Spec 85 Genes

44 Genes

Step 1 Step 2

+ Add a step

6. Examine the gene results. Do you see any genes that might be associated with parasite virulence?

View the gene pages of some of your results by clicking on the Gene ID. You can download the gene results by clicking on the button on the right.

L. infantum metacyclic protein expression *

Step 1: Mass Spec 246 Genes
Step 2: Mass Spec 44 Genes → 44 Genes

Organism Filter: Eubionidae (0) Trypanosomatida (44)

Gene Results (30 ortholog groups):

Gene ID	Transcript ID	Genomic Location (Gene)	Product Description
LINF_010008000	LINF_010008000-T1	LinJ.01:78,817..79,980(-)	acidocalcisomal exopolyphosphatase - putative
LINF_050005300	LINF_050005300-T1	LinJ.05:10,914..12,659(+)	paraffagellar rod component par4 - putative
LINF_060009000	LINF_060009000-T1	LinJ.06:129,917..131,056(+)	glutamine synthetase - putative
LINF_100010200	LINF_100010200-T1	LinJ.10:223,957..225,756(+)	GP63 - leishmanolysin
LINF_100010300	LINF_100010300-T1	LinJ.10:227,009..228,808(+)	GP63 - leishmanolysin
LINF_100010600	LINF_100010600-T1	LinJ.10:233,094..234,893(+)	GP63 - leishmanolysin
LINF_100010700	LINF_100010700-T1	LinJ.10:236,147..237,946(+)	GP63 - leishmanolysin
LINF_100010800	LINF_100010800-T1	LinJ.10:239,194..240,993(+)	GP63 - leishmanolysin
LINF_100010900	LINF_100010900-T1	LinJ.10:242,241..244,040(+)	GP63 - leishmanolysin
LINF_100011000	LINF_100011000-T1	LinJ.10:245,288..247,087(+)	GP63 - leishmanolysin
LINF_100011100	LINF_100011100-T1	LinJ.10:248,335..250,134(+)	GP63 - leishmanolysin
LINF_110005900	LINF_110005900-T1	LinJ.11:27,051..28,475(+)	seryl-tRNA synthetase - putative
LINF_1300013600	LINF_1300013600-T1	LinJ.13:297,855..299,441(-)	mitochondrial processing peptidase alpha subunit - putative
LINF_180005900	LINF_180005900-T1	LinJ.18:24,644..27,082(-)	alpha glucosidase II subunit - putative
LINF_180012600	LINF_180012600-T1	LinJ.18:304,871..306,271(+)	Elongation factor Tu - mitochondrial - putative
LINF_210013400	LINF_210013400-T1	LinJ.21:237,092..238,516(+)	proteasome regulatory non-ATPase subunit 5 - putative

Download

LINF_100010200

17 Proteomics

Mass Spec-based Expression Evidence

Mass Spec.-based expression evidence in Protein Browser (Data sets)

Proteomics data displayed as mapped peptides in the protein browser (JBrowse). Hover over or click a peptide for more details.

Transcript ID	Total Sequence Count	Unique Sequence Count	Sum Spectrum Count	Number of Samples	Protein Browser
LINF_100010200-T1	10	8	26	2	Interactive

View in protein browser

JBrowse viewer showing 125, 250, 375, 500 amino acids. Promastigote and amastigote proteomes (JPGCWA-67/TMAP-265) MS/MS Peptides Bretherston et al. Metacyclic Stage Proteome MS/MS Peptides Ouellette, et al. unpublished

Gene pages have a **Proteomics** section where you can view mapped peptides and data from other experiments.