



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.	✓

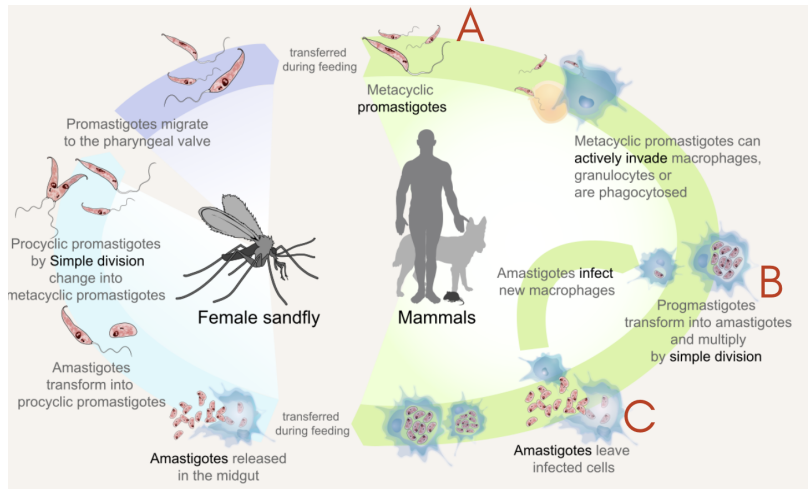
Introduction

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](https://www.tritrypdb.org). From the sidebar or header, search or scroll for “proteomics” and click on MassSpec. Evidence



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

infantum

Leishmania

Leishmania infantum JPCMS

Metacyclic stage (pH 5-6)

Post-translationally 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) (Tsiganik 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.

- Specifically, subtract the genes that have protein expression in the promastigote and amastigote stages.
- To do this, add a step to your strategy that combines the first step with other genes using a 1 minus 2 operator.
- 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

Combine with other Genes

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

proteome

Proteomics

Mass Spec. Evidence

Post-Translational Modification

Quantitative Mass Spec. Evidence

Use Genomic Colocation to combine with other features

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?

***L. infantum* metacyclic protein expression ***

Step 1: Mass Spec 162 Genes

Step 2: Mass Spec 985 Genes

66 Genes

+ Add a step

Add a step to your search strategy

Search for Genes by Mass Spec. Evidence

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

Configure Search Learn More View Data Sets Used

Reset values to default

Experiments and Samples

10 selected, out of 152

select only these | add these | clear these

inf

Leishmania

- ☒ **Leishmania infantum JPCMS**
 - ☐ Metacyclic Stage Proteome (Duellette, et al. unpublished)
 - ☐ metacyclic stage (pH 5-6)
 - ☐ Post-translationally 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) (Teigenkov et al.)**
 - ☒ amastigote phosphopeptides
 - ☒ promastigote phosphopeptides
 - ☒ **Promastigote and amastigote proteomes (MHOM/MA/67/ITMAP-243) (Briderton 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

Run Step

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.
 - a. Click on the edit button in the first step of the search
 - b. Click on the Revise option in the popup
 - c. Change the value of the "Minimum Number of Unique Peptide Sequences" search parameter from 1 to 5 and click on the revise button
 - d. Remember to do this for each step!

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

The image illustrates the process of increasing search stringency in a mass spectrometry search tool. It consists of four main components:

- Workflow Diagram (Top Left):** Shows a search process for *L. infantum* metacyclic protein expression. Step 1 is 'Mass Spec' (162 Genes) and Step 2 is 'Mass Spec' (985 Genes). Arrows indicate the flow from Step 1 to Step 2, with 'Edit' buttons for each step.
- Details for step Mass Spec (Top Right):** A popup window showing the configuration for Step 1. It includes a 'Revise' button (highlighted with a red arrow) and a table of search parameters:

Parameter	Value
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
- Configure Search Panel (Bottom Left):** A detailed view of the search configuration. It shows the 'Minimum Number of Unique Peptide Sequences' set to 5 (highlighted with a red arrow). The 'Apply min # peptide sequences / sample OR across samples' is set to 'Per Sample'. A 'Revise' button is circled at the bottom (highlighted with a red arrow).
- Revised Workflow Diagram (Bottom Right):** Shows the updated search process. Step 1 is 'Mass Spec' (85 Genes) and Step 2 is 'Mass Spec' (248 Genes). The 'Add a step' button is highlighted with a red arrow.

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 *

Mass Spec 44 Genes → Add a step

44 Genes (30 ortholog groups)

Some Genes in your combined result have Transcripts that were not returned

Gene Results | Genome View | Analyze Results

Show only the Genes in my basket.

Rows per page: 20

Download

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(+)	paraflagellar rod component part - 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.10:227,099..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_130013600	LINF_130013600-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-ATP-ase subunit 5 - putative

LINF_100010200

expand all | collapse all

Search section names...

- 1 Gene models
- 2 Annotation, curation and identifiers
- 3 Link outs
- 4 Genomic Location
- 5 Literature
- 6 Taxonomy
- 7 Orthology and synteny
- 8 Phenotype
- 9 Genetic variation
- 10 Transcriptomics
- 11 Sequences
- 12 Sequence analysis
- 13 Structure analysis
- 14 Protein features and properties
- 15 Function prediction
- 16 Pathways and interactions
- 17 Proteomics**
- 18 Immunology

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

Reset view

Reference sequence: 125 250 375 500

Peptides: LINF_100010200-T1-p1, LINF_100010200-T1-p1.599 (399 S)

Peptides: Promastigote and amastigote proteomes (RHON/NA/17/ITMAP-263) MS/MS Peptides Brotherton et al.

Peptides: 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.