

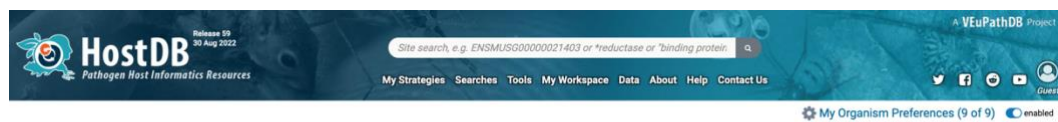
Host Response

Learning objectives:

- Exploring host responses by running a search strategy in HostDB
- Add steps in a search strategy
- Revising steps in a search strategy

1. Identify host genes that are differentially regulated in response to infection by different pathogens. For this exercise use <http://hostdb.org>

- The goal of this exercise is to find genes that are differentially regulated in the host during infection by multiple pathogens:
 - *Babesia rossi*
 - *Plasmodium vinckei*
 - *Toxoplasma gondii*
- Start by running search for genes that are differentially regulated (fold-change) in the experiment “*Babesia infection in dog (Smith et al. 2021)*”



Identify Genes based on RNA-Seq Evidence

Legend: S Similarity DE Differential Expression FC Fold Change P Percentile SA SenseAntisense		
Filter Data Sets: <input type="text"/> 38 rows		
Organism	Data Set	Choose a Search
<i>Bos taurus</i> breed Hereford	Transcriptome of <i>Bos taurus</i> during infection with virulent and avirulent <i>N. caninum</i> strains (Hercado et al.)	<input type="checkbox"/> DE <input type="checkbox"/> FC <input type="checkbox"/> P <input type="checkbox"/> SA
<i>Bos taurus</i> breed Hereford	Host cell transcriptome in bovine cells infected with <i>Cryptosporidium parvum</i> (Widmer et al.)	<input type="checkbox"/> DE <input type="checkbox"/> FC <input type="checkbox"/> P <input type="checkbox"/> SA
<i>Canis lupus familiaris</i> isolate SID07034	<i>Babesia</i> infection in dog (Smith et al. 2021) NEW	<input type="checkbox"/> DE <input checked="" type="checkbox"/> FC <input type="checkbox"/> P <input type="checkbox"/> SA
<i>Cavia porcellus</i> 2N	mRNA-based anti-tick (<i>I. scapularis</i>) vaccine (Sajid et al. 2021)	<input type="checkbox"/> FC <input type="checkbox"/> P <input type="checkbox"/> SA
<i>Gallus gallus</i> isolate bGalGal1	Chicken caecal tissue during <i>E. tenella</i> infection (Sandholt et al. 2021)	<input type="checkbox"/> DE <input type="checkbox"/> FC <input type="checkbox"/> P <input type="checkbox"/> SA
<i>Gallus gallus</i> isolate bGalGal1	Chicken macrophage-like cells (HD11) infected in vitro with <i>Eimeria tenella</i> (Sandholt et al. 2021) NEW	<input type="checkbox"/> DE <input type="checkbox"/> FC <input type="checkbox"/> P <input type="checkbox"/> SA
<i>Monodelphis domestica</i> BEE	Transcriptome of <i>Monodelphis domestica</i> stimulated by <i>Leishmania</i> (Shah et al. 2011)	<input type="checkbox"/> DE <input type="checkbox"/> FC <input type="checkbox"/> P <input type="checkbox"/> SA

- Configure the search to return all genes that are up or down regulated by at least 2-fold when comparing all low and high infection samples to Day 0 (low and high).

Identify Genes based on C. lupus familiaris isolate SID07034 Babesia infection in dog RNA-Seq (fold change)

Up or down regulated

A maximum of four samples are shown when more than four are selected.

For each gene, the search calculates:

$$\text{fold change}_{\text{up}} = \frac{\text{average expression value in comparison}}{\text{average expression value in reference}}$$

$$\text{fold change}_{\text{down}} = \frac{\text{average expression value in reference}}{\text{average expression value in comparison}}$$

and returns genes when $\text{fold change}_{\text{up}} \geq 2$ or $\text{fold change}_{\text{down}} \geq 2$.

You are searching for genes that are up or down regulated between at least two reference samples and at least two comparison samples.

- How many genes did the search return? What is the host organism?

1,732 Genes (1,580 ortholog groups)

Some Genes in your result have Transcripts that did not meet the search criteria. Explore

Gene Results | Genome View | Analyze Results

Genes: 1,732 Transcripts: 4,147 Show Only One Transcript Per Gene

Gene ID	Transcript ID	Organism	Product Description
ENSGAF00845007235	ENSGAF008450127	Canis lupus familiaris isolate SID07034	cathelicidin antimicrobial peptide [Source:NCBI gene (formerly Entrezgene) Acc:442947]

- Add a step and navigate to the RNA-Seq datasets

Add a step to your search strategy

Combine with other Genes

Transform into related records

Use Genomic Collocation to combine with other features

Choose how to combine with Genes

Choose which Genes to combine. From...

Find genes based on their expression levels quantified by RNA-Seq (RNAseq or RNA sequencing)

- Find the experiment "Mouse transcriptome during P. vinckei CY infection time course (Ramaprasad et al)" and select the fold-change search.

Search for Genes by RNA-Seq Evidence

The results will be  intersected with  the results of Step 1.

Legend: S Similarity DE Differential Expression FC Fold Change P Percentile SA SenseAntisense

Filter Data Sets: 2 results (filtered from a total of 38)

Organism ?	Data Set	Choose a Search
<i>Bos taurus</i> breed Hereford	Host cell transcriptome in bovine cells infected with <i>Cryptosporidium parvum</i> (Widmer et al.)	DE FC P SA
<i>Mus musculus</i> C57BL6J	Mouse transcriptome during <i>P. vinckei</i> CY infection time course (Ramaprasad et al)	DE FC P SA

- Configure the search to return all genes that are differentially regulated by at least 2-fold when comparing the 18hr and 24hr infection to the 6 and 12 hr infection samples.

Differential Expression **Fold Change** Percentile SenseAntisense

For the Experiment
☒ Mouse transcriptome during *P. vinckei* CY infection time course - Sense
☐ Mouse transcriptome during *P. vinckei* CY infection time course - Antisense

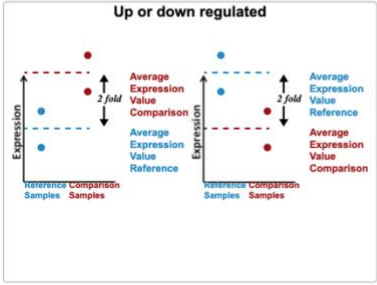
return Genes
 that are
 with a **Fold change** \geq
 between each gene's expression value
 (or a Floor of)
 in the following **Reference Samples**
☒ 6 hr
☒ 12 hr
☐ 18 hr
☐ 24 hr

and its expression value
 (or the Floor selected above)
 in the following **Comparison Samples**
☐ 6 hr
☒ 12 hr
☒ 18 hr
☒ 24 hr

SeqEvidence#GenesByRNASeqmmusC57BL6J_Ramaprasad_PvinCY_IDC_ebi_rnaSeq_RSRC

Example showing one gene that would meet search criteria
 (Dots represent this gene's expression values for selected samples)


Up or down regulated



For each gene, the search calculates:

$$\text{fold change}_{\text{up}} = \frac{\text{average expression value in comparison}}{\text{average expression value in reference}}$$

$$\text{fold change}_{\text{down}} = \frac{\text{average expression value in reference}}{\text{average expression value in comparison}}$$
 and returns genes when $\text{fold change}_{\text{up}} \geq 2$ or $\text{fold change}_{\text{down}} \geq 2$.
 You are searching for genes that are **up or down regulated** between at least two **reference samples** and at least two **comparison samples**.

 COMMUNITY CHAT



- How many results did you get?

- Is this result of 0 genes surprising? What can you do to change this? Hint: think about the host organism for each experiment, are they the same or different? What can you do to compare across species?



Add a step to your search strategy ?

Your Genes from Step 1 will be converted into Orthologs

Organism

1 selected, out of 9

select only these | add these | clear these

mus

- ☒ Mammalia
 - ☒ Rodentia
 - ☒ Muridae
 - ☒ Mus musculus C57BL6J [Reference]

select only these | add these | clear these

Syntenic Orthologs Only?

no

Run Step

- Use the same logic above and add a step to identify all genes that are differentially regulated in the experiment “Mouse transcriptome during early and late chronic infection with T. gondii (Garfoot et al.)”

- Configure the search to compare the 21- and 28-day infected female mice to the 21- and 28-day uninfected female mice.

← Add a step to your search strategy ⓘ

Differential Expression **Fold Change** Percentile

For the Experiment
 ⓘ Mouse transcriptome during early and late chronic infection with T. gondii unstranded
 ⓘ return protein coding ⓘ Genes
 ⓘ that are up or down regulated ⓘ
 ⓘ with a Fold change >= 2 ⓘ
 ⓘ between each gene's average ⓘ expression value ⓘ
 ⓘ (or a Floor of 10 reads ⓘ)
 ⓘ in the following Reference Samples ⓘ

☐ 21 days Tg Female
☒ 21 days Uninfected Female
☐ 28 days Tg Female
☒ 28 days Uninfected Female
☐ 3 mo Tg Female
☐ 3 mo Uninfected Female
 select all | clear all

and its average ⓘ expression value ⓘ
 ⓘ (or the Floor selected above)
 ⓘ in the following Comparison Samples ⓘ

☒ 21 days Tg Female
☐ 21 days Uninfected Female
☒ 28 days Tg Female
☐ 28 days Uninfected Female
☐ 3 mo Tg Female
☐ 3 mo Uninfected Female
 select all | clear all

Example showing one gene that would meet search criteria
 (Dots represent this gene's expression values for selected samples)

Up or down regulated

For each gene, the search calculates:

$$\text{fold change}_{\text{up}} = \frac{\text{average expression value in comparison}}{\text{average expression value in reference}}$$

$$\text{fold change}_{\text{down}} = \frac{\text{average expression value in reference}}{\text{average expression value in comparison}}$$

and returns genes when $\text{fold change}_{\text{up}} \geq 2$ or $\text{fold change}_{\text{down}} \geq 2$.

You are searching for genes that are up or down regulated between at least two reference samples and at least two comparison samples.

- How many genes did you get? Are there any gene with interesting functions in the list of results? Hint: Try a GO enrichment

Opened (1) All (3) Public (2) Help

Unnamed Search Strategy ⓘ

Step 1: Chip Beasies infection RNAseq... 1,732 Genes
 Step 2: Orthologs... 2,824 Genes
 Step 3: 188 Genes
 Step 4: 87 Genes

87 Genes (76 ortholog groups)

Some Genes in your combined result have Transcripts that were not returned by one or both of the two input searches. Explore

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

Genes: 87 Transcripts: 268 Show Only One Transcript Per Gene

Gene ID	Transcript ID	Genomic Location (Gene)	Product Description
ENSMUSG000000070034	ENSMUST0000000935...	mmusC57BL6J_chr1:85,576,899..85,598,817()	Sp110 nuclear body protein [Source:MGJ SymbolAcc:MGJ:1923364]
ENSMUSG000000070034	ENSMUST0000001283...	mmusC57BL6J_chr1:85,576,899..85,598,817()	Sp110 nuclear body protein [Source:MGJ SymbolAcc:MGJ:1923364]
ENSMUSG000000070034	ENSMUST0000001319...	mmusC57BL6J_chr1:85,576,899..85,598,817()	Sp110 nuclear body protein [Source:MGJ SymbolAcc:MGJ:1923364]

87 Genes (76 ortholog groups)

Some Genes in your combined result have Transcripts that were not returned by one or both of the two input searches. Explore

Gene Results | Genome View | **Gene Ontology Enrichment**

Find Gene Ontology terms that are enriched in your gene result. View More

Parameters

Analysis Results:

GO ID	GO Term	Genes in the input with this term	Genes in your result with this term	Percent of input genes in your result	Fold enrichment	Odds ratio	P-value	Benjamini	Bonferroni
GO:0097367	carbohydrate derivative binding	2248	33	1.5	3.69	5.41	7.66e-12	1.68e-9	2.45e-9
GO:0043148	anion binding	2817	34	1.4	3.54	5.23	1.05e-11	1.68e-9	3.35e-9
GO:0035639	purine ribonucleoside triphosphate binding	1841	29	1.6	3.96	5.52	4.32e-11	4.67e-9	1.38e-8
GO:1901265	nucleoside phosphate binding	2179	31	1.4	3.58	5.07	9.37e-11	5.91e-9	3.06e-8
GO:0004164	nucleotide binding	2179	31	1.4	3.58	5.07	9.37e-11	5.91e-9	3.06e-8
GO:0032555	purine ribonucleoside binding	1920	29	1.5	3.80	5.27	1.18e-10	5.91e-9	3.79e-8
GO:0017076	purine nucleotide binding	1924	29	1.5	3.77	5.22	1.41e-10	5.91e-9	4.30e-8
GO:0032553	ribonucleotide binding	1938	29	1.5	3.77	5.21	1.48e-10	5.91e-9	4.73e-8