EuPathDB RNAseq BigWig & expression value export & integration tool

Note: using EuPathDB *My Datasets* and the EuPathDB Galaxy workspace requires you to log in to the site since the data is stored privately under your account.

Currently this tool works on workflows that use Cufflinks to generate expression values. It is recommended that all data files used are grouped using the collections method in Galaxy. For more information about collections please see the Galaxy Collections help pages. An example workflow is available on the EuPathDB Galaxy front page for you to use and try out.

Once a workflow has run successfully you can export and integrate your data privately into EuPathDB by following these steps:

1. Click on the "EuPathDB Export Tools" menu item in the left-hand panel, then click on the "RNAseq to EuPathDB" export tool.

2. The RNAseq export tool parameters will appear in the middle section of the Galaxy page. Fill out the parameters completely:
   a. Enter a meaniful name for your dataset (you can always modify this).
   b. Select the BigWig collection from the list.
   c. Select the Cufflinks **gene** expression collection.
   d. Enter a summary for your dataset (you can always modify this).

e. Enter a description for your dataset (you can always modify this).
f. Click on the "Execute" button.

3. After you click on the execute button a new step is initiated in the right hand history pannel: ExportToEuPathDBinfo.html. The step will initially be grey, then yellow when running and green when completed. This should not take a long time.

4. Once the export is complete, go to the EuPathDB site where your data will be visualized. For example, if your data is *Plasmodium berghei* RNAseq data go to PlasmoDB.org

5. Click on the "My Datasets" link in the grey menu bar.

6. The "My Datasets" page includes a list of all datasets that you have exported to EuPathDB with the most recent export at the top of the list.

7. Click on the experiment name in the Name/ID column to enter that dataset's page. If the experiment is still being installed into EuPathDB you should notice the status line indicating that the dataset is currently being installed.

8. Once the dataset is installed, it is ready for use (note the status line indicates it is ready with a green checkmark). You can access the results in two ways:
   a. Access available searches against your data
   b. Access GBrowse tracks representing your data

My Dataset: Experiment_A1 ✏

Status: ✅ This data set is installed and ready for use in ToxoDB.

Owner: Me

Description: samples from experiment A1 ✏

ID: 12993

Data Type: **RNASeq** (RnaSeq 1.0)

Summary: Experiment A1 ✏

Created: an hour ago

Dataset Size: 295.20 M

Quota Usage: 3.09% of 10.00 G

Available Searches: • genes by RNA-Seq user dataset (fold change)

Use This Dataset in ToxoDB

🧩 Compatibility Information ❓

EuPathDB Website	Required Resource	Required Resource Release	Installed Resource Release
ToxoDB	TgondiiME49 Genome	29	29

This dataset is compatible with the current release, build 40, of **ToxoDB**. It is installed for use.

📊 GBrowse Tracks

Filename	GBrowse Status	
day3_rep2.bw	○ This file has not been added to **GBrowse**.	Send To **GBrowse** ⬆
day3_rep1.bw	○ This file has not been added to **GBrowse**.	Send To **GBrowse** ⬆
Display a menu	○ This file has not been added to **GBrowse**.	Send To **GBrowse** ⬆

9. To run a search against your data, click on the "genes by RNA-Seq user dataset (fold change)" link.

10. The next page provides you with customizable search parameters with you analyzed samples available for selection.

11. Once you configure the search parameters, you can click on the "Get Answer" button to get your results.

12. The results page includes custom columns based on your data including: fold change, expression values for the reference and comparator (FPKM) and a profile graph. The results can be further intersected with results from other searches in EuPathDB or analyzed using one of the available analysis tools such as GO enrichment.

13. The data is also represented on gene pages in the Transcriptomics section under the heading "User Dataset Transcriptomics Graphs" .

Synonymous SNPs All Strains 4

Total SNPs All Strains 28

10 Transcriptomics

▸ Transcript Expression ☰ Data sets

▾ User Dataset Transcriptomics Graphs

Search this table... 🔍   Showing 2 rows

▸	↕ Preview	↕ Name	↕ Assay Type
▾		**Experiment_A1**	**RNA-seq**

TGME49_263310 - UserDataset 12993

Full Dataset Description

▸ Data table

X-axis
Stage/Sample

Y-axis
Transcript levels of fragments per kilobase of exon model per million mapped reads (FPKM).

Choose gene for which to display graph
🔘 TGME49_263310

Choose graph(s) to display
☑ fpkm

Graph options
☐ Show log Scale (not applicable for log(ratio) graphs, percentile graphs or data tables)

14. To view RNAseq coverage graphs in the genome browser, go back to the "My Datasets" section, select the dataset you want to view and scroll down to the GBrowse Tracks section.

Data Type: **RNASeq** (RnaSeq 1.0)
Summary: Experiment A1 ✏
Created: 2 hours ago
Dataset Size: 295.20 M
Quota Usage: 3.09% of 10.00 G
Available Searches: • genes by RNA-Seq user dataset (fold change)

Use This Dataset in ToxoDB

🧩 Compatibility Information ❓

EuPathDB Website	Required Resource	Required Resource Release	Installed Resource Release
ToxoDB	TgondiiME49 Genome	29	29

This dataset is compatible with the current release, build 40, of **ToxoDB**. It is installed for use.

📊 GBrowse Tracks

Filename	GBrowse Status	
day3_rep2.bw	⭘ This file has not been added to **GBrowse**.	Send To **GBrowse** ⬆
day3_rep1.bw	⭘ This file has not been added to **GBrowse**.	Send To **GBrowse** ⬆
day4_rep1.bw	⭘ This file has not been added to **GBrowse**.	Send To **GBrowse** ⬆
day4_rep2.bw	⭘ This file has not been added to **GBrowse**.	Send To **GBrowse** ⬆

15. Click on the "Send To GBrowse" buttons.

🧩 Compatibility Information ❓

EuPathDB Website	Required Resource	Required Resource Release	Installed Resource Release
ToxoDB	TgondiiME49 Genome	29	29

This dataset is compatible with the current release, build 40, of **ToxoDB**. It is installed for use.

📊 GBrowse Tracks

Filename	GBrowse Status	
day3_rep2.bw	✅ Sent to GBrowse.	View In **GBrowse** ❯
day3_rep1.bw	✅ Sent to GBrowse.	View In **GBrowse** ❯
day4_rep1.bw	⭕ Sending to GBrowse...	Send To **GBrowse** ⬆
day4_rep2.bw	⭕ Sending to GBrowse...	Send To **GBrowse** ⬆

16. Now you can click on the "View in GBrowse" button to visulaize your data in the genome browser.