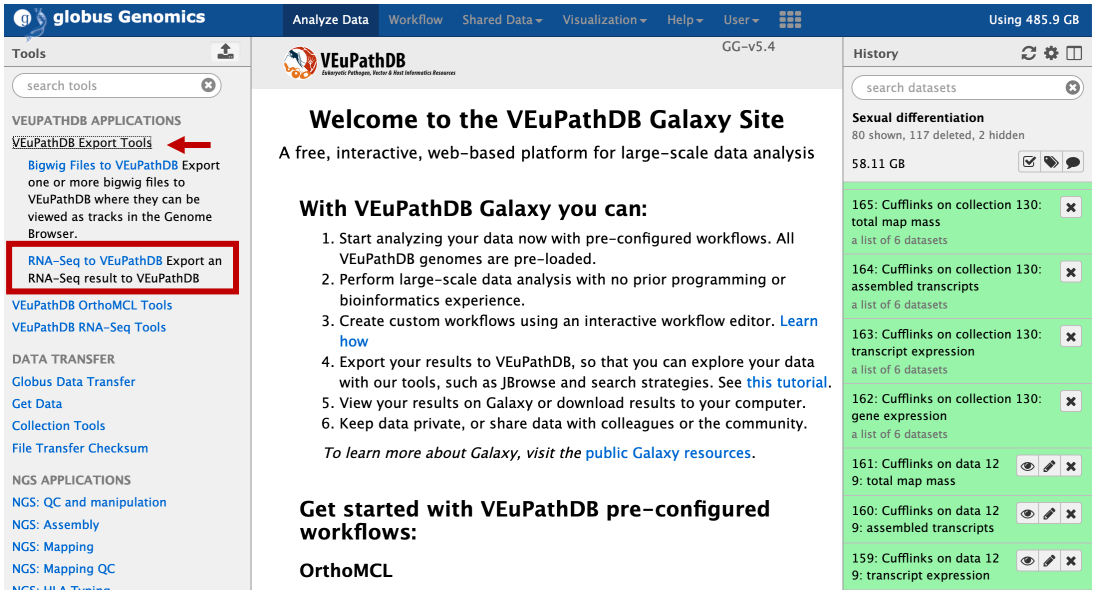
VEuPathDB RNA-Seq BigWig & expression value export & integration tool

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

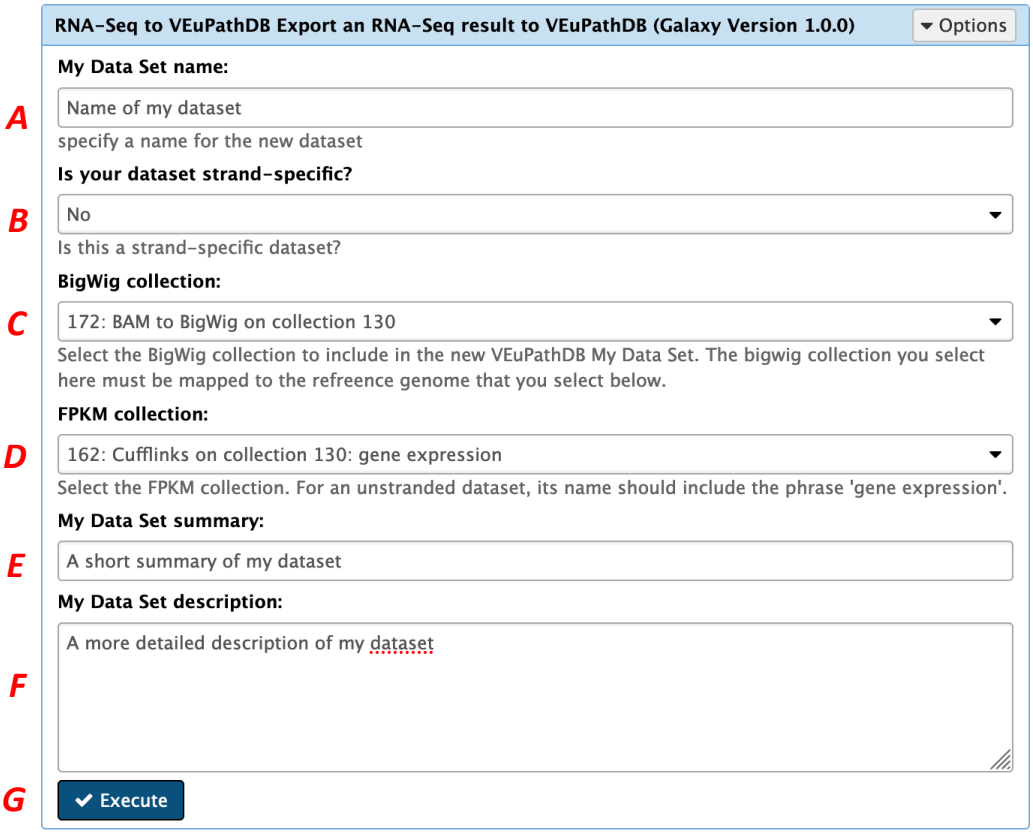
Currently this tool works on workflows that use Cufflinks to generate expression values. It is recommended that all data files used are grouped using the collections method in Galaxy. For more information about collections please see the Galaxy Collections help pages. An example workflow is available on the VEuPathDB 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 VEuPathDB by following these steps:

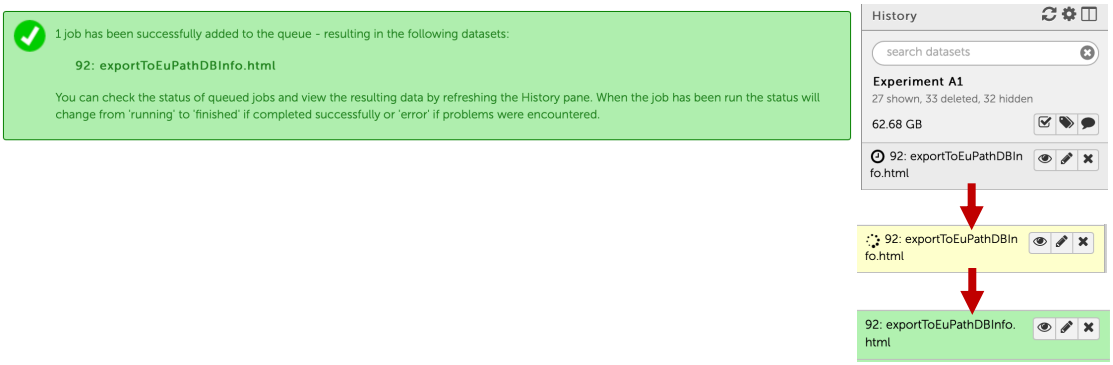
1. Click on the "VEuPathDB Export Tools" menu item in the left-hand panel, then click on the "RNA-Seq to VEuPathDB" export tool.



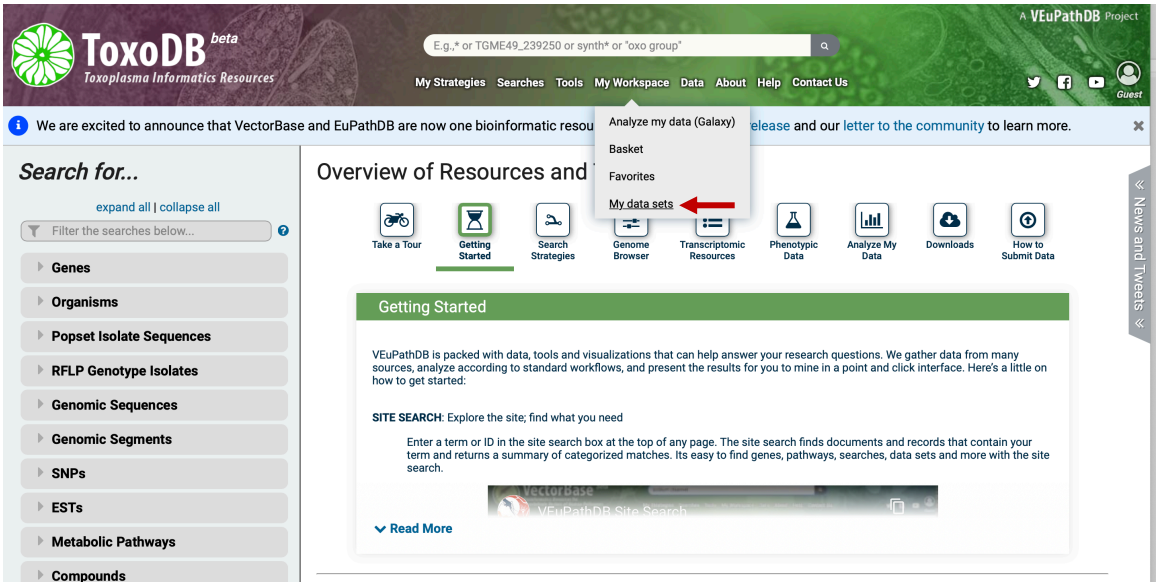
2. The RNA-Seq export tool parameters will appear in the middle section of the Galaxy page. Fill out the parameters completely:
    a. Enter a meanigful name for your dataset (you can always modify this).
    b. Select whether you RNA-Seq dataset is strand specific or not.
    c. Select the BigWig collection from the list.
    d. Select the Cufflinks **gene** expression collection.
    e. Enter a summary for your dataset (you can always modify this).
    f. Enter a description for your dataset (you can always modify this).
    g. Click on the “Execute” button.



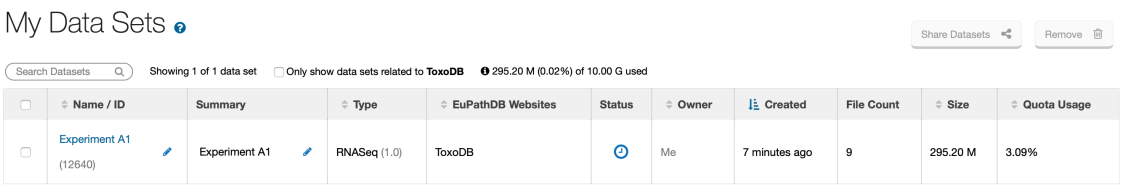
3. After you click on the execute button a new step is initiated in the right hand history pannel: ExportToVEuPathDBinfo.html. The step will initially be grey, then yellow when running and green when completed. This should not take a long time.



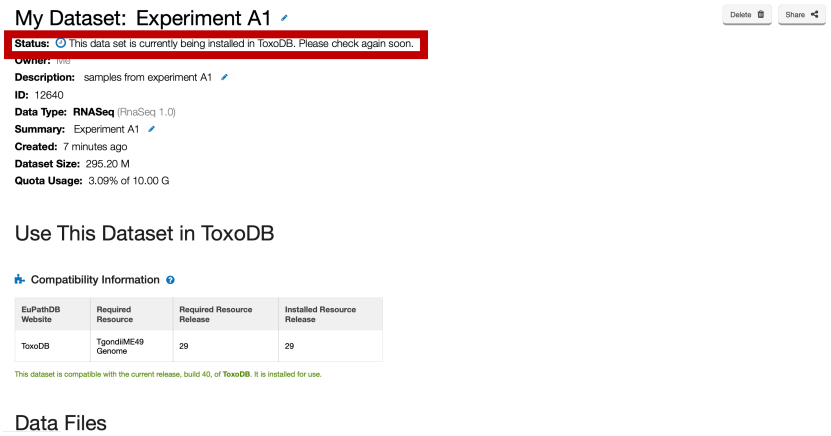
4. Once the export is complete, go to the VEuPathDB site where your data will be visualized. For example, if your data is *Plasmodium berghei* RNA-Seq data go to PlasmoDB.org
5. Click on the “My Workspace” menu in the header then select the *My data* sets link.



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

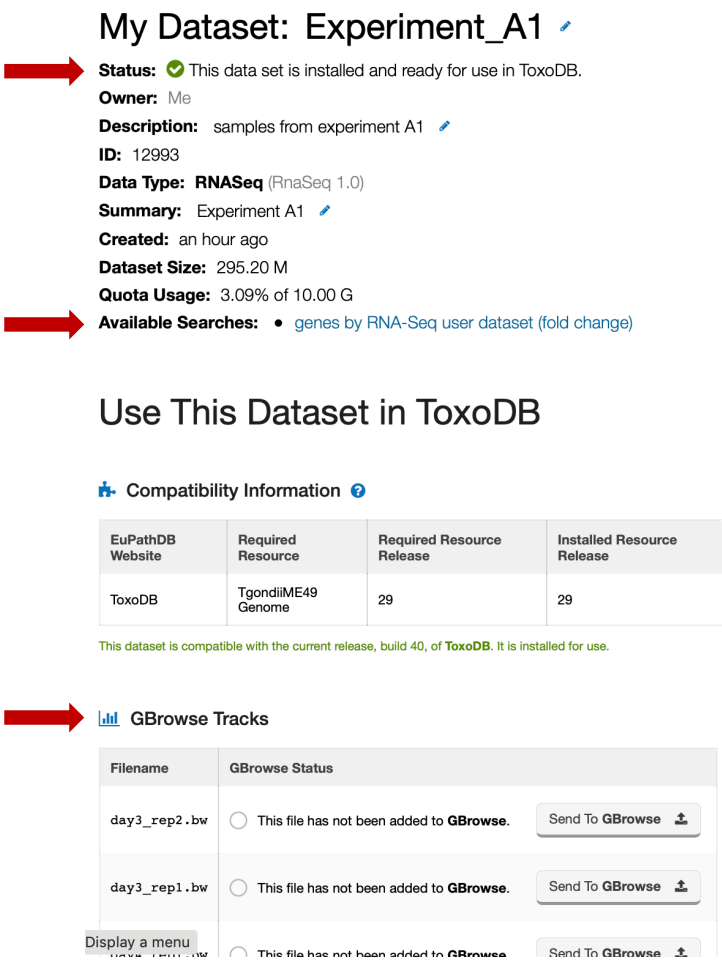


7. Click on the experiment name in the Name/ID column to enter that dataset’s page. If the experiment is still being installed into VEuPathDB you should notice the status line indicating that the dataset is currently being installed.



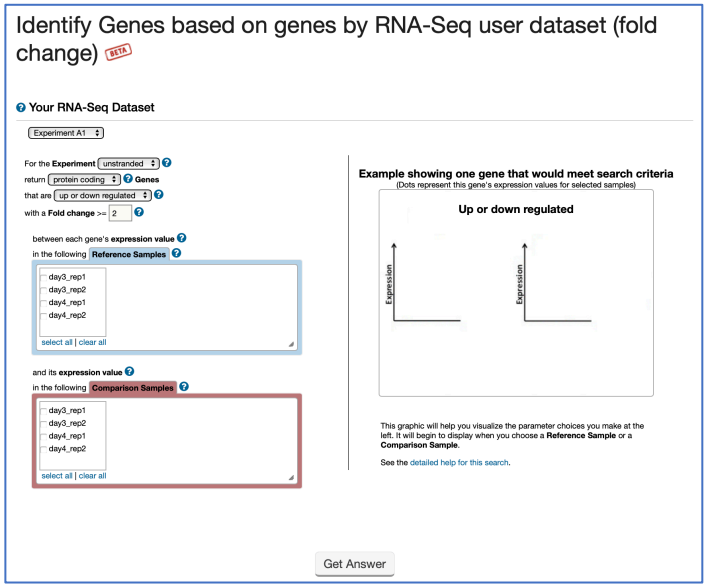
8. Once the dataset is installed, it is ready for use (note the status line indicates it is ready with a green checkmark). You can access the results in two ways:

a. Access available searches against your data
b. Access GBrowse tracks representing your data

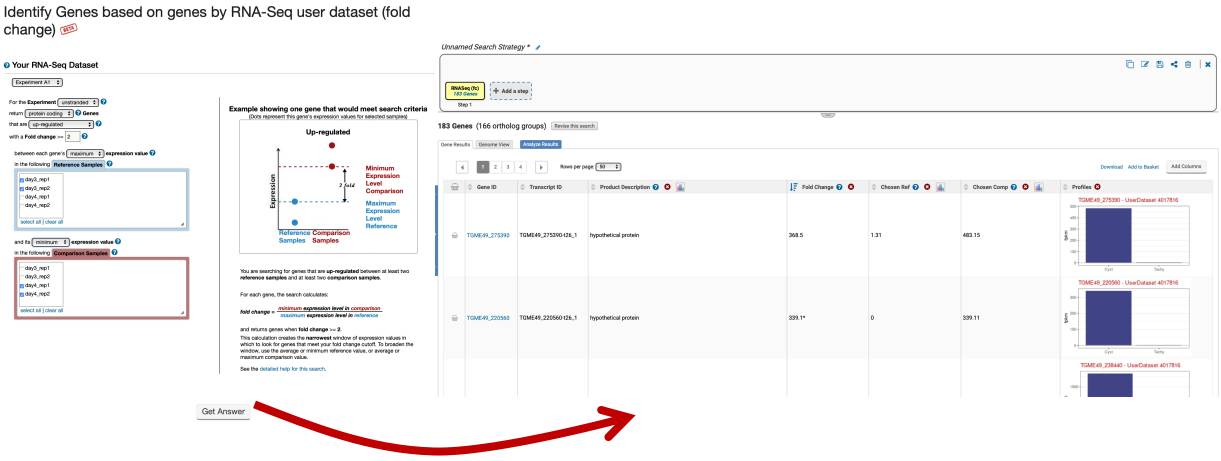


9. To run a search against your data, click on the “genes by RNA-Seq user dataset (fold change)” link.

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

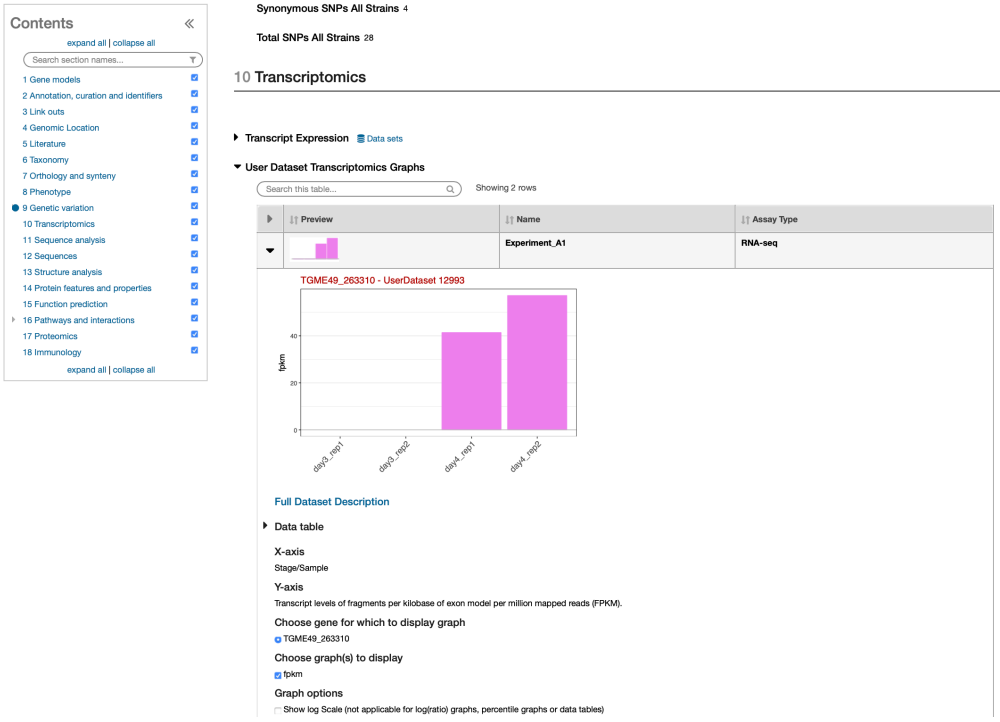


11. Once you configure the search parameters, you can click on the “Get Answer” button to get your results.

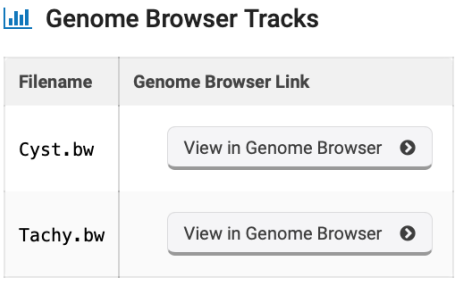


12. The results page includes custom columns based on your data including: fold change, expression values for the reference and comparator (FPKM) and a profile graph. The results can be further intersected with results from other searches in VEuPathDB or analyzed using one of the available analysis tools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

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



14. To view RNA-Seq 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 Genome Browser Tracks section.



15. Now you can click on the “View in Genome Browser” button to visulaize your data in the genome browser.