WASHINGTON STATE UNIVERSITY

Tripal Galaxy

Execution of Scientific Workflows for Tripal-Based Community Databases

Shawna Spoor, Connor Wytko, Ming Chen, Abdullah Almsaeed, Bradford Condon, Heidi Hough, Nic Herndon, Nick Mills, Margaret Staton, Jill Wegrzyn, Alex Feltus, Dorrie Main, Stephen P. Ficklin

Galaxy

Galaxy is an open source online platform for scientific data integration and analysis.

- Makes computational biology accessible to everyone
- Allows for multi-step computational analyses
  - simple linear process
  - browser based user interface

Tripal Galaxy

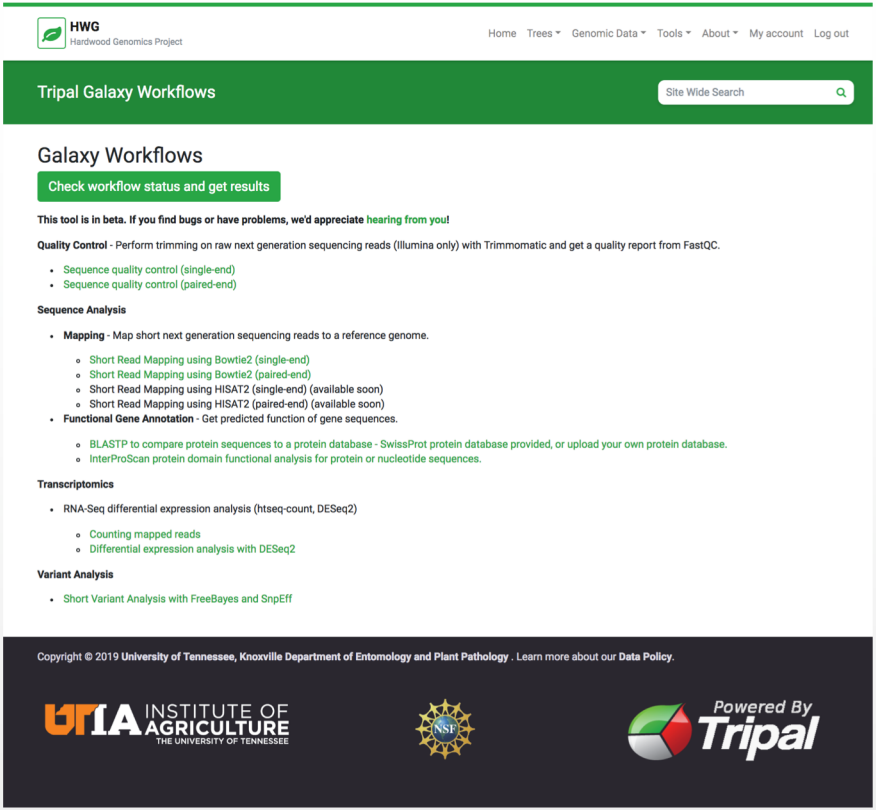
The Tripal Galaxy module allows Tripal-based sites to use the Galaxy project. That means:

- Analytical tools for users within your website
  - Consistent user interface
  - The power of high performance computing on your website

Tripal Galaxy: Using the Module

The module can be used in two ways:

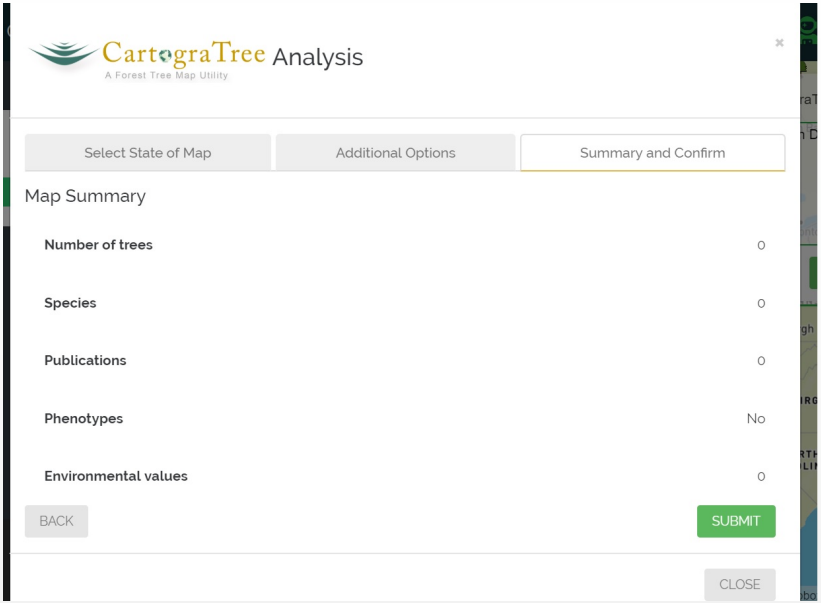
- Using a graphical user interface



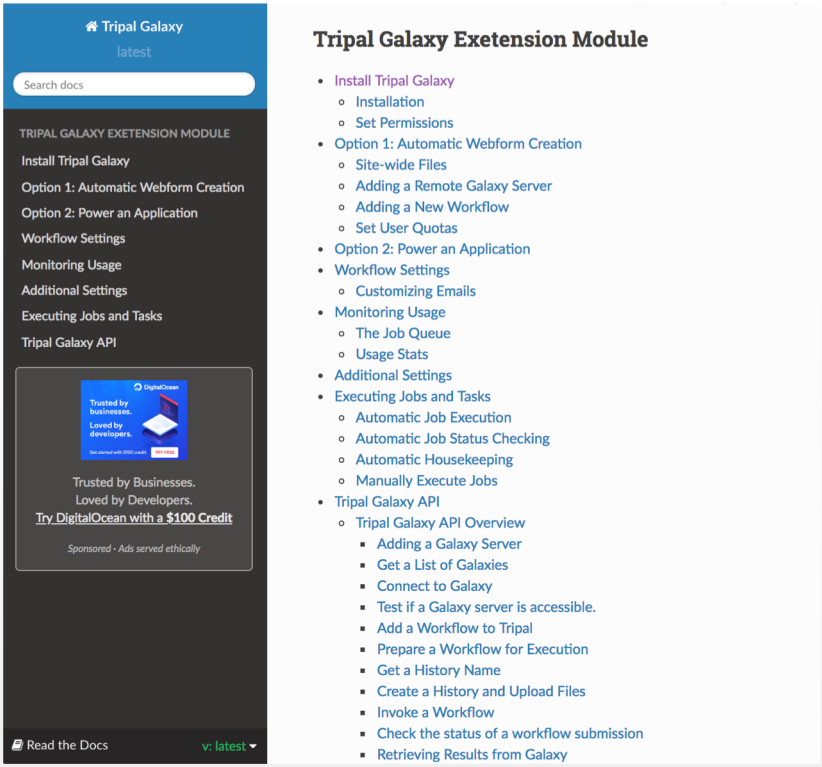
Tripal Galaxy: Using the Module

The module can be used in two ways:

- Through the application programming interface (API)

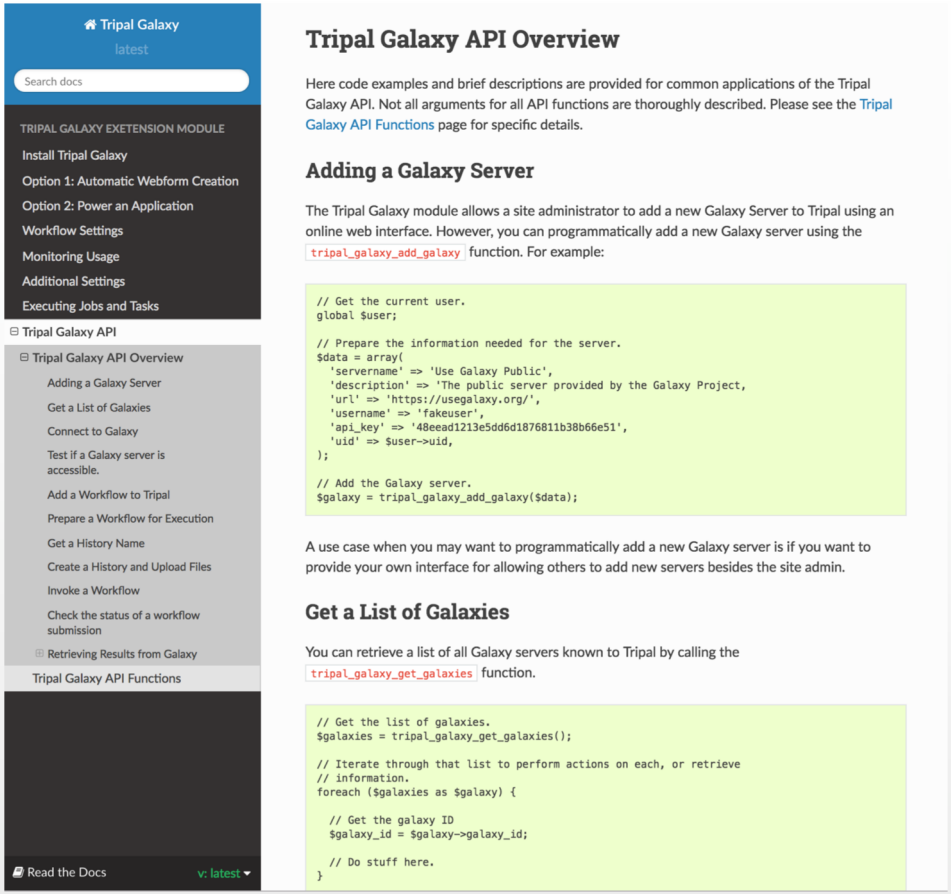


Tripal-Galaxy Docs



https://tripal-galaxy.readthedocs.io/en/latest/

Tripal-Galaxy Docs



Installing Tripal Galaxy: blend4php

The blend4php library was developed as part of the Tripal Gateway Project as stand-alone PHP bindings to the Galaxy RESTful API. It allows any PHP-based web application to communicate with a remote Galaxy instance.

https://github.com/galaxyproject/blend4php

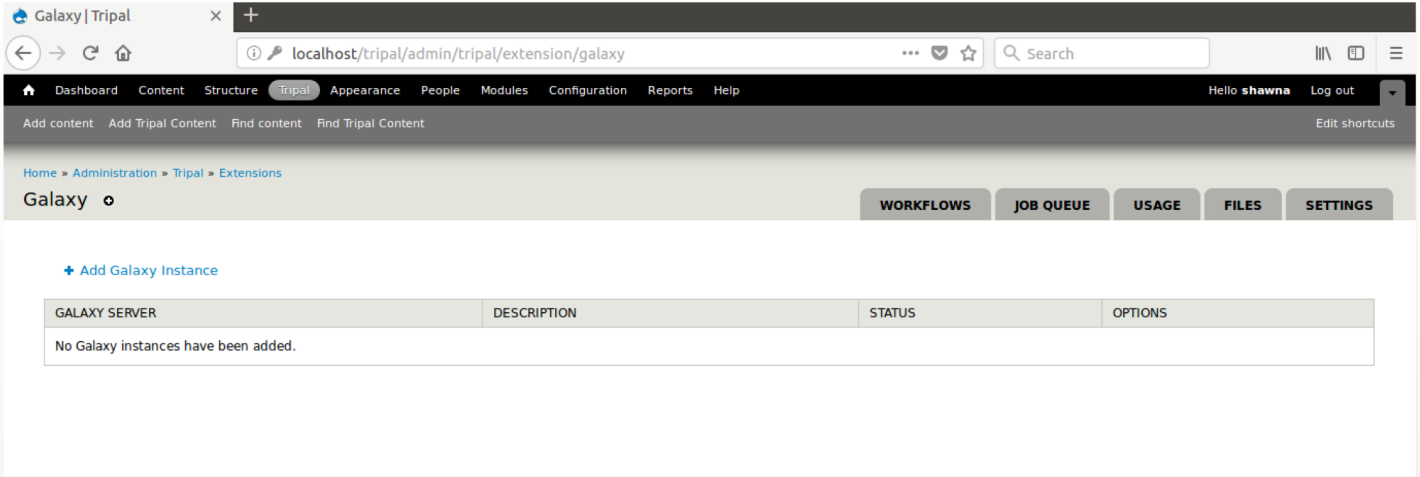
Installing Tripal Galaxy: module installation

Module installation can be done using:

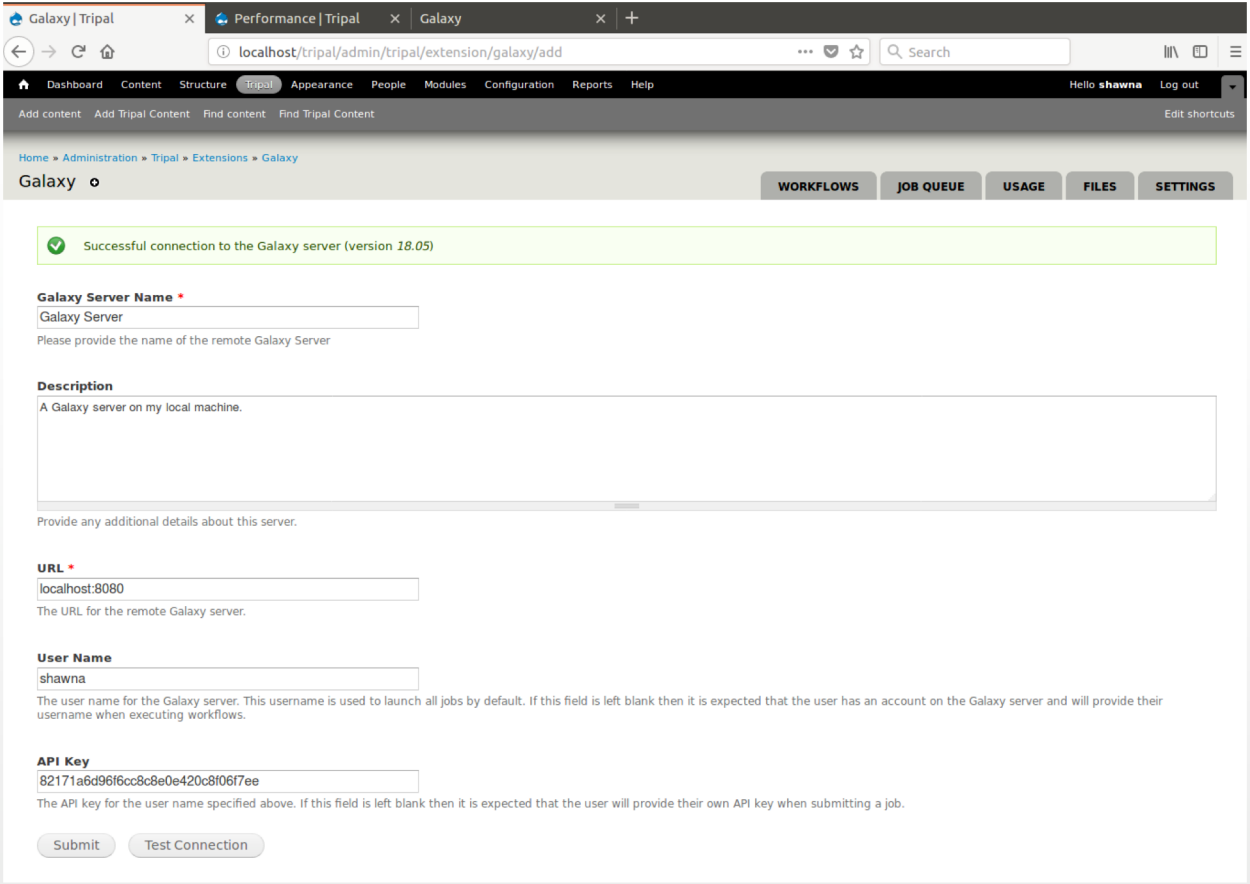
- Drush: `drush pm-enable tripal_galaxy`
- GUI: admin/modules/install
- Git clone:

```
git clone
git@github.com:tripal/tripal_galaxy.git
```

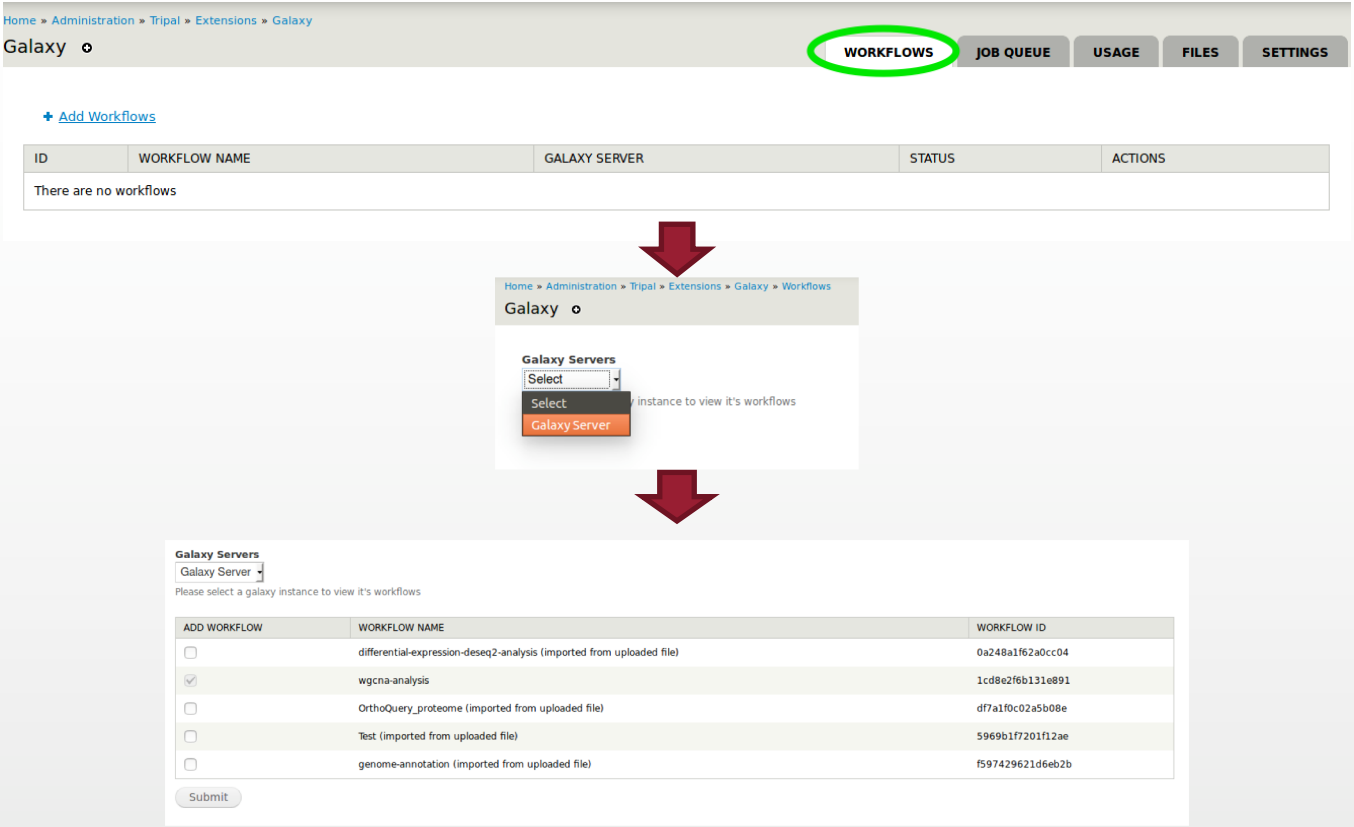
Using Tripal Galaxy: Setting it up



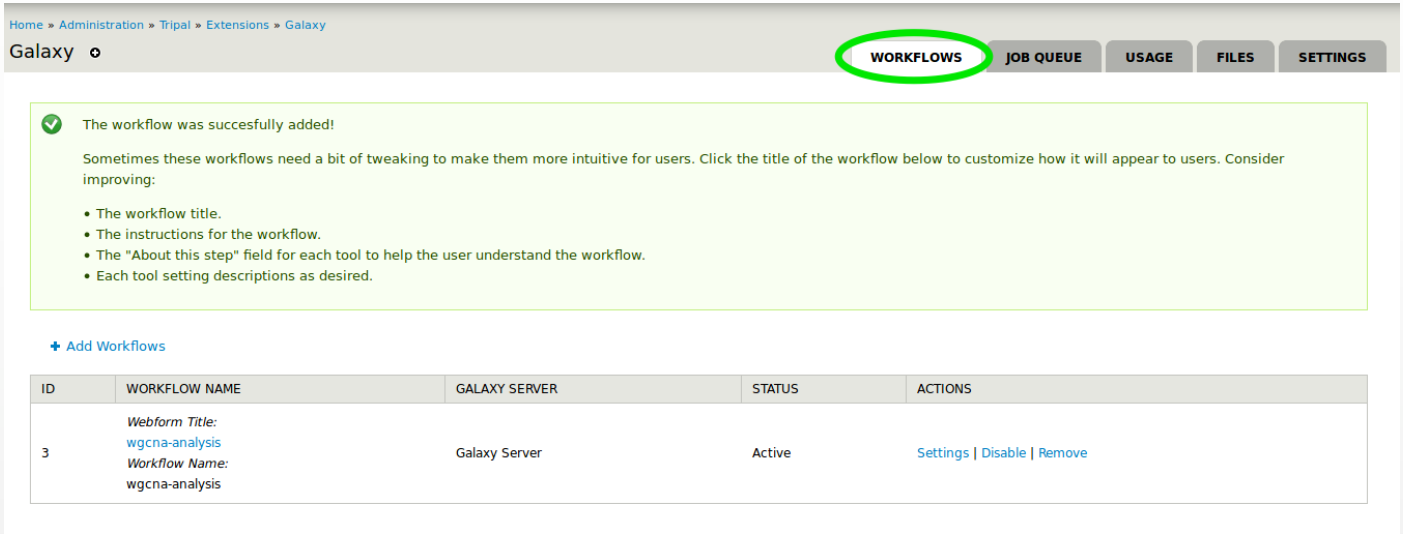
Using Tripal Galaxy: Adding a Galaxy Instance



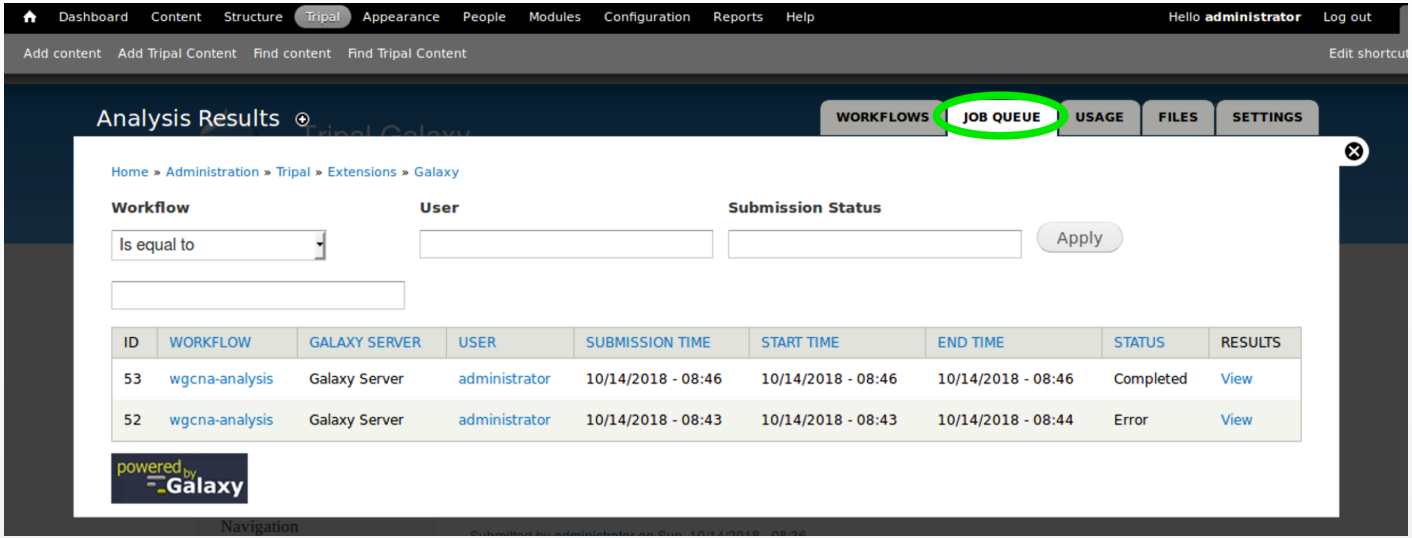
Using Tripal Galaxy: Adding a Workflow



Using Tripal Galaxy: Adding a Workflow



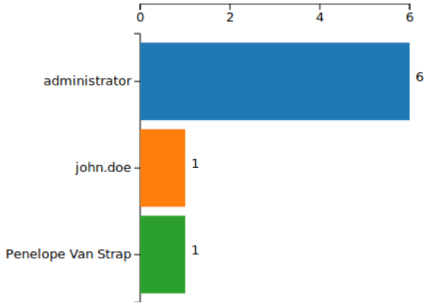
Using Tripal Galaxy: Job Queue



Using Tripal Galaxy: Usage Stats

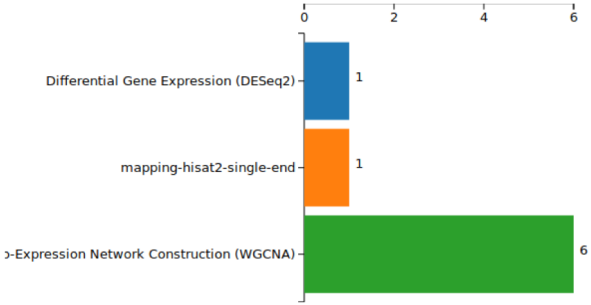
Top 10 Submitters

The following bar chart shows the top 10 users that have submitted workflows.



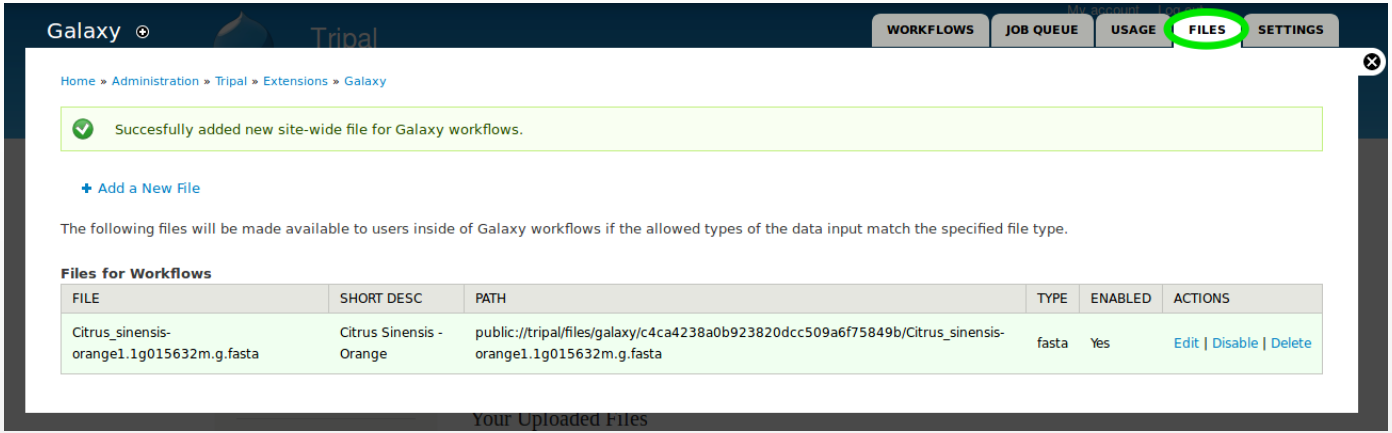
Top 10 Workflows Submitted

The following bar chart shows the top 10 workflows that have been submitted.



Using Tripal Galaxy: Files

Upload files and make them available to all users.



Using Tripal Galaxy: Running a Workflow

Co-Expression Network Construction (WGCNA)

Step 1 — Step 2 — Step 3 — Step 4 — Step 5 — Step 6 — Preview — Submit

Page 1 of 8 (0%)

Step 1: Instructions

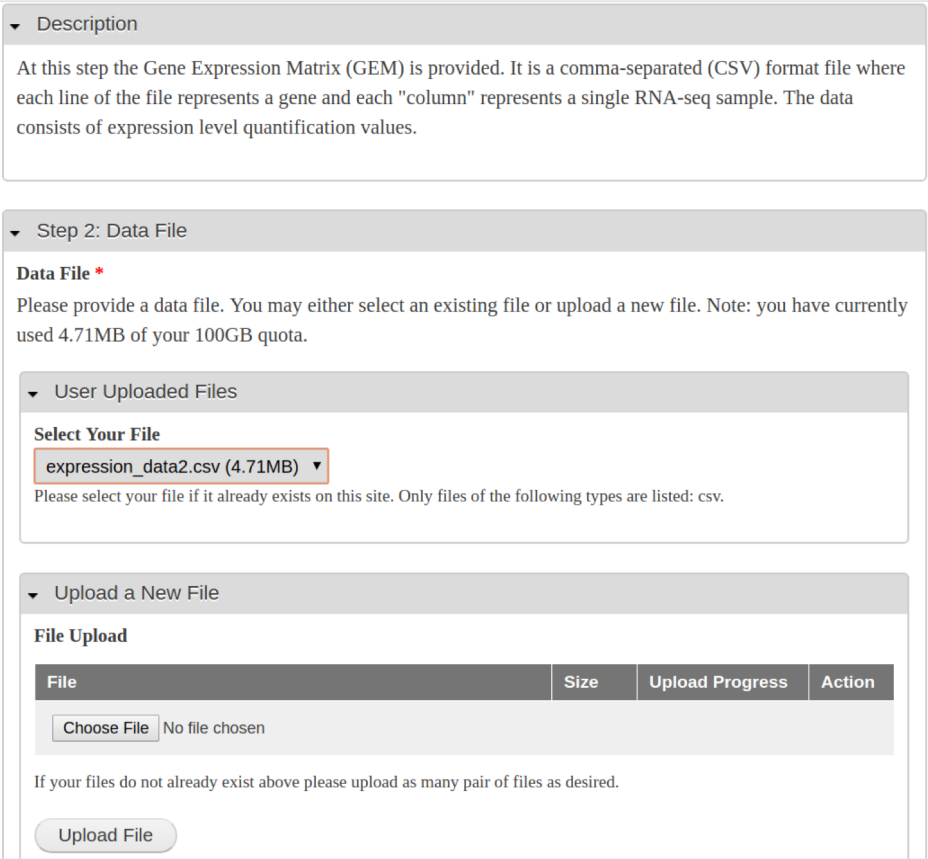
Gene Co-Expression Networks GCNs are often used to help identify groups of genes that are co-functional. The analysis requires a file containing a Gene Expression Matrix (GEM) containg gene expression levels derived from either RNA-seq or microarray data sets. Genes that tend to be expressed at similar levels across a set of samples are considered co-expressed. A correlation value (often Pearson's or Spearman's correlation) is assigned to each gene pair. WGNCA is an R-based package that performs pair-wise gene correlations and uses a method referred to as soft-thresholding to cull pair-wise correlation values that are non signficant. To learn more about WGCNA and co-expression network construction please see the WGCNA homepage.

EditExecution of this workflow is supported by Galaxy.

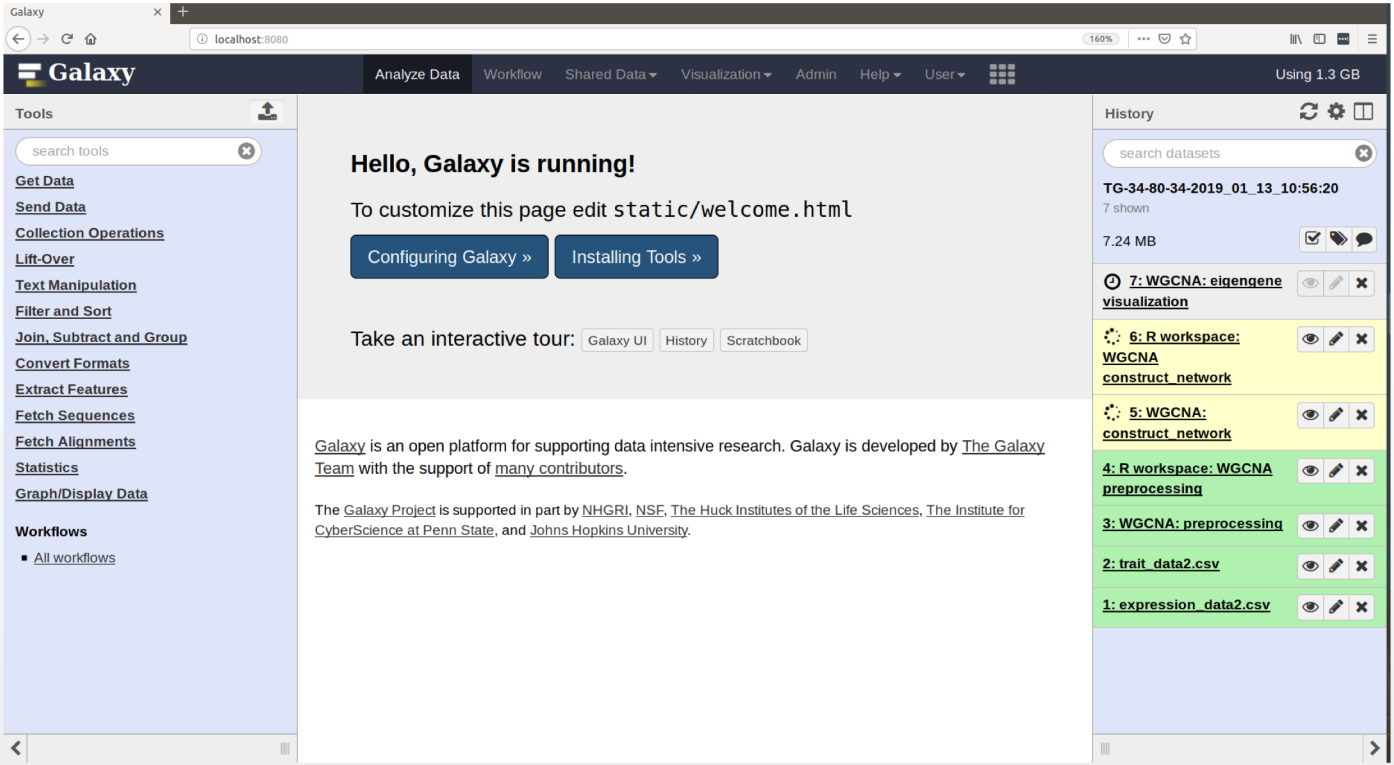
This workflow uses the following tools:

- **WGCNA: preprocessing (version 1.0.2)**: Data clearning and preprocessing.
- **WGCNA: construct network (version 1.0.2)**: Construct gene network.
- **WGCNA: eigengene visualization (version 1.0.2)**: Eigengene visualization.

Using Tripal Galaxy: Running a Workflow



Using Tripal Galaxy: Running a Workflow



Using Tripal Galaxy: Running a Workflow

Analysis Results

Analysis Name
Co-Expression Network Construction (WGCNA)

▸ Submission Details

Results
All files associated with the analysis, including input files and all output files are listed below. If an error occurred during creation of any file it will be noted. Expand the field for the result below to view the status, peek or download result files.

- **1: expression_data2.csv (ok)**
- **2: trait_data2.csv (ok)**
- **3: WGCNA: preprocessing (ok)**
- **4: R workspace: WGCNA preprocessing (ok)**
- **5: WGCNA: construct_network (ok)**
- **6: R workspace: WGCNA construct_network (ok)**
- **7: WGCNA: eigengene visualization (ok)**

WGCNA: eigengene visualization

Download Results
WGCNA: eigengene visualization (1.43MB)

View Results
WGCNA: eigengene visualization
Note: It may take a few moments for the results page to load.

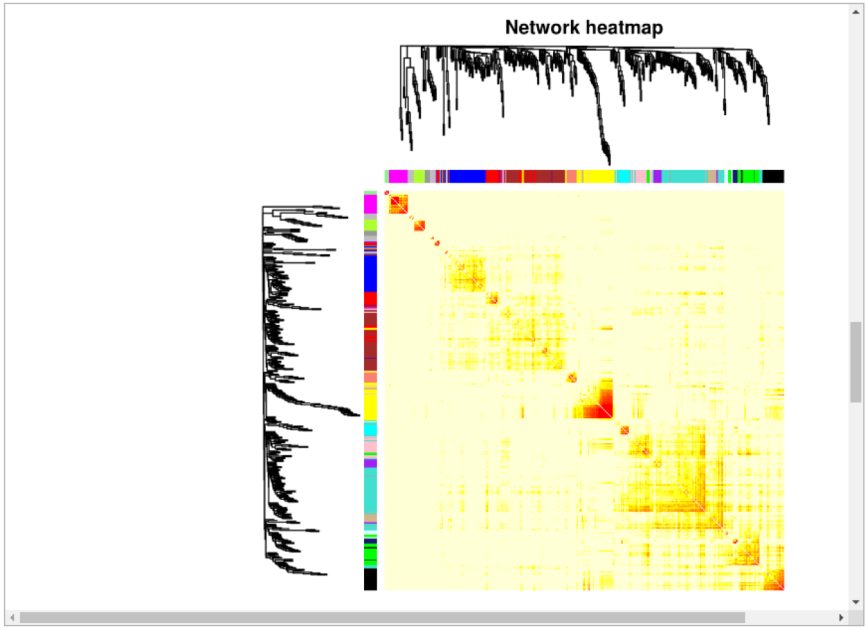
Using Tripal Galaxy: Running a Workflow

Home » administrator » Analyses » Analysis Results

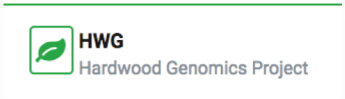
Results Viewer

Viewing results for WGCNA: eigengene visualization

View full screen



Tripal Galaxy in use: Hardwood Genomics



https://www.hardwoodgenomics.org/

https://github.com/statonlab/galaxy-workflows

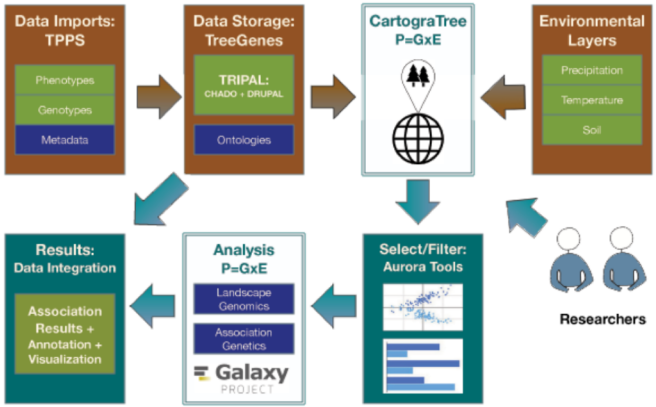
https://github.com/statonlab/aurora-galaxy-tools

Tripal Galaxy in use: Cartogratree

CartograTree About User Guide Data Events TreeGenes

CartograTree

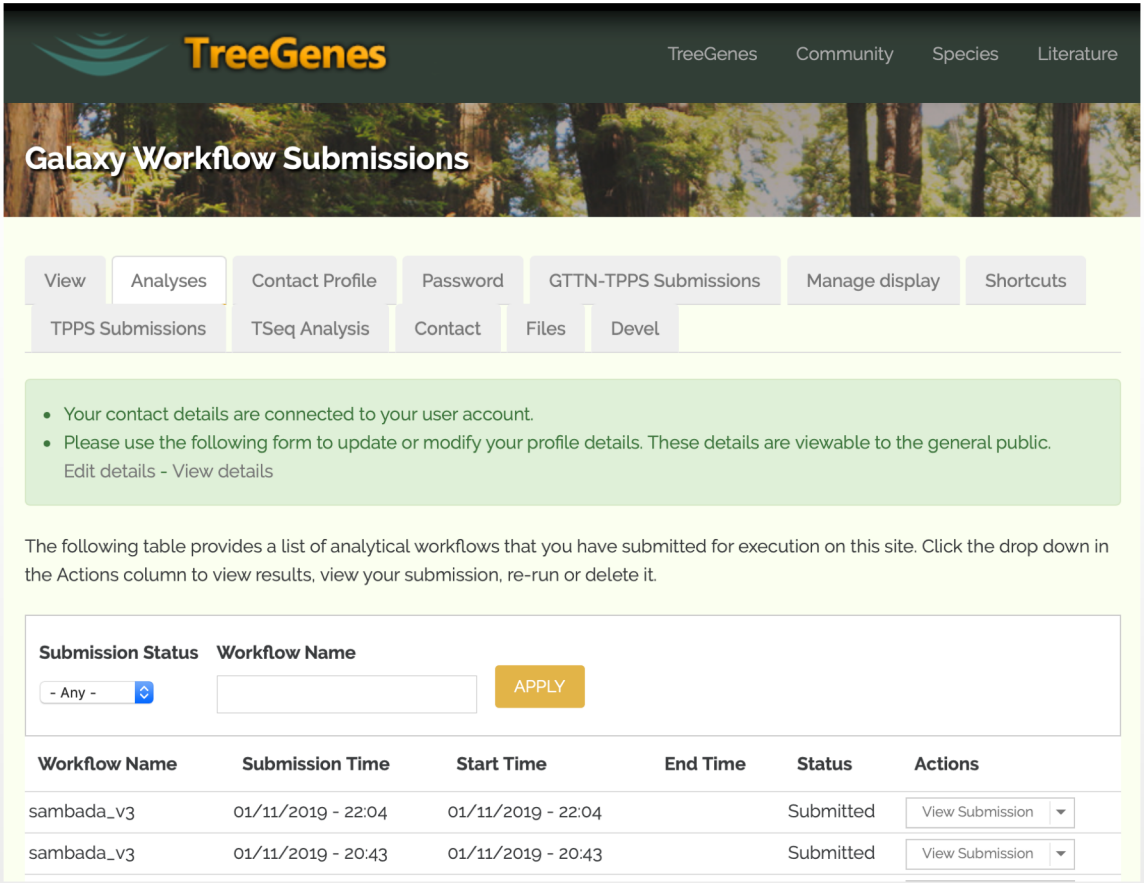
Integration of genomic, phenomic, and environmental data for georeferenced forest trees



Forest trees are long-lived and immobile individuals that serve as ideal models to assess population structure and adaptation to the environment. Despite the availability of comprehensive data, the researchers who study them are challenged to integrate data describing genotype, phenotype, and the environment. Towards this goal, the web application CartograTree was designed and implemented as an open repository and open-source analytic web-based framework for all three.

Learn More

Tripal Galaxy in use: Cartogratree



The Tripal Galaxy Team

Washington State University

Stephen Ficklin's Team
Stephen Ficklin
Connor Wytko
Shawna Spoor

Dorrie Main's Team
Dorrie Main
Heidi Hough

The University of Tennessee

Meg Staton's Team
Meg Staton
Ming Cheng
Abdullah Almsaeed
Bradford Condon

Clemson University

Alex Feltus Team
Alex Feltus
Nick Mills

University of Connecticut

Jill Wegrzyn's Team
Jill Wegrzyn
Nic Herndon

Acknowledgements

NSF Award #1443040

Further Information:

Tripal Galaxy Docs:
http://tripal.info/tutorials/v3.x/galaxy

Tripal Galaxy github:
https://github.com/tripal/tripal_galaxy

Galaxy:
https://galaxyproject.org/

Tripal:
http://tripal.info/

Hardwood Genomics:
https://www.hardwoodgenomics.org/

TreeGenes:
https://treegenesdb.org