HPRC Maroon Galaxy
What is the Galaxy Project?

- public Galaxy instance
- many popular bioinformatic tools are available
- no programming knowledge required
- accessed using a web browser
- reproducible workflow
- shared data and workflows
- try usegalaxy.org to see if Galaxy is a good fit for your project

The Galaxy Project is supported in part by NSF, NHGRI, The Huck Institutes of the Life Sciences, The Institute for CyberScience at Penn State, and Johns Hopkins University.
There are many tutorials available with example input data and step by step analysis for various topics.
### Galaxy Shared Workflows

- can be used to share your workflow used in a publication
- input files are not stored here

#### Published Workflows

<table>
<thead>
<tr>
<th>Name</th>
<th>Annotation</th>
<th>Owner</th>
<th>Community Rating</th>
<th>Community Tags</th>
<th>Last Updated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic RNA-Seq Analysis - Differential Expression</td>
<td>From the RNA-Seq analysis tutorial during the Functional Genomics Workshop 2012 <a href="https://caps.osu.edu/pgw-workshop">https://caps.osu.edu/pgw-workshop</a></td>
<td>mejia-guerra</td>
<td>⭐⭐⭐⭐⭐</td>
<td></td>
<td>Jun 22, 2012</td>
</tr>
<tr>
<td>RNA-seq differential analysis (single-end short reads, 2 conditions, 2 replicates)</td>
<td>Workflow based on Tophat and cuffdiff. Inputs: 4 fastq files (experiments), 1 bam file (pseudoreads), 1 gtf file (annotations). Outputs: bam, bamtwig, xis,...</td>
<td>rna-seq-hellinger-group</td>
<td>⭐⭐⭐⭐⭐</td>
<td>illumina, rna-seq, cuffdiff</td>
<td>Jul 17, 2013</td>
</tr>
<tr>
<td>RNA-seq differential expression analysis</td>
<td>RNA-seq differential analysis</td>
<td>chmy</td>
<td>⭐⭐⭐⭐⭐</td>
<td>easy</td>
<td>Aug 11, 2017</td>
</tr>
</tbody>
</table>

search results for rna-seq and sorted by Community Rating
Galaxy Shared Workflow Example
Galaxy Tool Shed  
https://toolshed.g2.bx.psu.edu

- tool repositories are created by the Galaxy staff as well as Software developers and Galaxy users
- software versioning is available within the toolshed
- more than just bioinformatics software tools are available
  - statistics
  - graphing
  - machine learning
The Galaxy Project has Community Support for Feature Requests and Resolving Issues
HPRC Maroon Galaxy on the Ada Cluster
HPRC Maroon Galaxy on the Ada Cluster

- TAMU HPRC Galaxy instance
- Accessed using a web browser
- Has been available to TAMU students, staff and faculty since 2015
  - will be retired on June 30, 2021
  - no more jobs submitted after May 31, 2021
HPRC Maroon Galaxy on the Ada Cluster

- Installed in 2015 (v15.07) as a joint effort by Ping Luo (HPRC) and Dr. Rodolfo Aramayo (Department of Biology) - (currently 75 registered users)
- There are 350+ users across all HPRC Galaxies that are used for teaching and research
- Auto-installation of tools did not consistently work
- No support for compressed files (gzip) other than decompression on upload
- No auto-detection of fastq/sanger format on upload
  - user had to run FastQC to evaluate quality score encoding or run Fastq Groomer
- Most newer software tools require v16.01+
- Early support for running tools on a cluster did not work well so tools were hard coded to use a specific amount of resource (cores, memory, time)
  - multiple tool configurations had to be added to support different resource requirements
HPRC Maroon Galaxy on the Grace Cluster
HPRC Maroon Galaxy on the Grace Cluster

- TAMU HPRC Galaxy instance
- Available now to TAMU students, staff and faculty

https://galaxy-grace.hprc.tamu.edu/maroon
HPRC Grace Maroon Galaxy Features

- Installed in Spring 2021 (v21.01) on the HPRC Grace cluster.
- This is a significant deployment for the HPRC bio community for research and teaching.
- Improved integration of automatically installing tools from the Galaxy Toolshed.
- New support added to maintain file compression after uploading
  - Some software doesn't support compressed file format as input but file decompression is supported
- Multiple tool versions are available within a single tool
- New auto-detection of fastqsanger format on upload
  - No need to use Fastq Groomer unless working with very early Illumina sequence data
- Integrated support for selecting cluster resources for each tool (cores, memory, time)
Galaxy Tutorials and Chat Features

Welcome to Galaxy Training!
Collection of tutorials developed and maintained by the worldwide Galaxy community

Galaxy for Scientists
<table>
<thead>
<tr>
<th>Topic</th>
<th>Tutorials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to Galaxy Analyses</td>
<td>10</td>
</tr>
<tr>
<td>Assembly</td>
<td>9</td>
</tr>
<tr>
<td>Climate</td>
<td>4</td>
</tr>
<tr>
<td>Computational chemistry</td>
<td>6</td>
</tr>
<tr>
<td>Ecology</td>
<td>6</td>
</tr>
</tbody>
</table>

Galaxy Tips & Tricks
<table>
<thead>
<tr>
<th>Topic</th>
<th>Tutorials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using Galaxy and Managing your Data</td>
<td>17</td>
</tr>
</tbody>
</table>

Galaxy for Developers and Admins
<table>
<thead>
<tr>
<th>Topic</th>
<th>Tutorials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Galaxy Server administration</td>
<td>10</td>
</tr>
</tbody>
</table>

The same tutorials found at galaxyproject.org/learn can be accessed here and followed without leaving Maroon Galaxy

More than 300 Galaxy users and admins discussing anything Galaxy related
Circos Tutorial Demonstration

Visualisation
Training material for data visualisation in Galaxy

Requirements
Before diving into this topic, we recommend you to have a look at:
- Introduction to Galaxy Analyses

Material

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Slides</th>
<th>Hands-on</th>
<th>Input dataset</th>
<th>Workflows</th>
<th>Galaxy tour</th>
<th>Galaxy instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visualisations in Galaxy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genomic Data Visualisation with JBrowse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visualization with Circos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Check Your HPRC SUs Balance

You can also change your default HPRC project account and view SUs charged for each job run during the current fiscal year.
Shared Data Libraries

Files can be added to a ‘Data Library’ which you can share with your Galaxy colleagues. Send a request to the HPRC helpdesk if you would like a Data Library for your group.
NCBI Blast Databases Available

Additional NCBI BLAST and custom databases can be added upon request.
HPRC Maroon Galaxy Access

- Try Galaxy at usegalaxy.org to see if it appropriate for your project
- How to access HPRC Maroon Galaxy on the Grace Cluster
  - Available to Texas A&M students, staff and faculty with a NetID and an HPRC account
  - Apply for an HPRC account first
    - https://hprc.tamu.edu/apply
  - Then send an email request for a Maroon Galaxy account
    - help@hprc.tamu.edu
  - Need to use TAMU VPN when connecting to Galaxy from off campus
  - Login to Maroon Galaxy using your TAMU NetID and password
- Read the Galaxy Usage Notes
  - https://hprc.tamu.edu/wiki/SW:Galaxy
Thank you.

Any questions?